

Informant Reports in Clinical Assessment

Andres De Los Reyes & Bridget A. Makol

Department of Psychology, University of Maryland at College Park, Comprehensive Assessment and Intervention Program, Department of Psychology, Biology-Psychology Building, Room 3123H, College Park, Maryland 20742, United States

Author Note

Efforts by the authors were supported by a grant from the Institute of Education Sciences (R324A180032).

Correspondence regarding this manuscript should be addressed to Andres De Los Reyes, Comprehensive Assessment and Intervention Program, Department of Psychology, University of Maryland at College Park, Biology/Psychology Building, Room 3123H, College Park, MD 20742; Office: 301-405-7049; Fax: 301-314-9566; E-mail: adlr@umd.edu; Twitter: @JCCAP_Editor

CITATION

De Los Reyes, A., & Makol, B.A. (2021). Informant reports in clinical assessment.

In G. Asmundson (Ed.), *Comprehensive clinical psychology*. (2nd ed.). Elsevier.

(in press)

Abstract

Clients display considerable variations in functioning across the contexts that encompass their social environments (e.g., home, school/workplace, peer interactions). No single measurement method can fully capture these variations. Yet, assessors must balance the need to accurately capture clients' clinical presentations, and at the same time attend to clinical feasibility when considering which measures to use. We describe one method—informants' subjective reports about clients' mental health—and highlight research and theory that connects scores on informants' reports with scores on non-subjective methods (e.g., observed behavior, performance-based tasks, official records). We also discuss issues in use and interpretation of informants' reports, provide recommendations for clinical assessments, and outline important directions for future research and practice.

Keywords: agreement; checklists; compensating operations; converging operations; correspondence; disagreement; diverging operations; informant discrepancies; multiple informants; Operations Triad Model; rating scales; surveys

1.0 Introduction*

Clients lead complex lives. Indeed, consider the lived experiences of individuals who experience mental health concerns. Where do adults feel anxious or sad? In what settings do children experience hyperactivity, aggression, or inattention? Under what circumstances do adolescents experience bullying, or the mental health consequences of bullying? Perhaps the logical answer to these questions is that these experiences manifest everywhere. Yet, do they manifest the same way, regardless of the demands of the environments that clients navigate? When clients experience mental health concerns, they likely vary in not only *what* brings them to a mental health professional, but also *where* or what contexts tend to elicit their challenges. After all, if we know one thing about mental health concerns it is this: they result from a complex interplay among psychological, socio-cultural, and biological factors that pose risk for, or protect against, maladaptive reactions to social contexts (e.g., Cicchetti, 1984; Luthar et al., 2000; Sanislow et al., 2010). However, not all contexts elicit maladaptive reactions to the same degree (e.g., Kazdin & Kagan, 1994; Mischel & Shoda, 1995). Therefore, individuals may display mental health concerns in some contexts, such as work or school settings, to a greater degree than other contexts, such as home. In fact, these contextual variations occur within a variety of mental health conditions, such as attention-deficit/hyperactivity disorder (ADHD), autism spectrum disorder (ASD), conduct problems, social anxiety, substance use, and social competence (e.g., Bögels et al., 2010; Deros et al., 2018; Dirks et al., 2012; Drabick et al., 2007, 2008; Kraemer et al., 2003; Lerner et al., 2017).

The complexity underlying clients' clinical presentations leads to a logical conclusion: We cannot capture all of what brings clients to the attention of a professional with a single index or number. Life is far too complicated for that. At the same time, researchers and practitioners have limited time and resources when conducting assessments to capture all of the complexity in clients' mental health. Regardless of the reason for carrying out an assessment, the general rule is to administer just enough instruments to balance out multiple considerations: (a) meeting clients' needs, (b) ensuring accuracy, and (c) maximizing feasibility. In this chapter, we provide an overview of informants' subjective reports about clients' mental health. Research indicates this method yields quick, accurate indices of clients' concerns, and facilitates both detecting needs for services and monitoring responses to services. We highlight challenges with using this method,

review the latest research and theory on how to make the most of this method in research and practice settings, and outline recommendations for future research and practice.

It deserves mention that much of the work reviewed in this chapter focuses on use of informant reports when assessing youth. Key aspects of the literature necessitate this focus. That is, developers of informant-based instruments in youth mental health research commonly create parallel versions of the instrument that are tailored to informants who have unique knowledge or *expertise* about the behavior of the youth undergoing evaluation (e.g., parents who rate child's home behavior, teachers who rate child's school behavior). The bodies of research and theory we discuss below capitalize on this core feature of youth assessment: the ability to collect reports about the same client's functioning from different kinds of informants, and not only compare data from these reports to each other but also against data from independent assessment modalities (e.g., observed behavior, task performance). In this chapter, we refer to this approach to assessment—examining patterns of multi-informant assessments in reference to independent assessment modalities—as the *comprehensive assessment* approach. While the literature on this approach is extensive and well-characterized in youth assessments, our understanding of how these kinds of assessments operate when assessing adults remains relatively nascent. Thus, the conceptual and empirical foundations of work discussed in this chapter draw heavily from research on youth assessments. That said, where possible we discuss these issues, highlight examples from the adult literature, and provide recommendations for future work that extends this research to adult assessment.

2.0 Who Do We Ask to Provide Reports?

Assessors decide which informants provide reports to collect, and which reports (if any) inform key decisions regarding service delivery, such as making diagnoses and planning treatment. In turn, when understanding which informants traditionally provide reports in mental health assessments, it is important to start with some history. Much of this history points to circumstances in which researchers and practitioners alike have *excluded* certain informants from the assessment process. To this day, some researchers and practitioners still exclude these informants' reports from the assessment processes. We raise these issues to highlight a key reality of mental health assessments: assessors vary in the degree to which they base decisions about which informants they include in assessments on sound science and defensible standards. Many times their decisions appear to rely more so on tradition and, indeed, conjecture, and when this occurs, we can often point to science that not only fails to corroborate these decision-making

strategies, but also supports alternative strategies. Given the complexity of assessing clients' concerns and the lack of empirically derived guidelines for integrating information sources, it is unsurprising that this phenomenon occurs (see also Beidas et al., 2015).

Consider two examples. First, for decades assessors of adult clients limited their information sources to the self-reports of those initiating or receiving care (Hunsley & Mash, 2007), perhaps based on the notion that soliciting reports from collateral informants like spouses or co-workers either lacked feasibility or resulted in information of little-to-no incremental value (Achenbach, 2006; De Los Reyes et al., 2019a). The data suggest otherwise. That is, not only has research demonstrated clinically feasible strategies for gathering reports from informants other than the adult clients seeking care, but also that these informants provide incrementally valuable information, over-and-above adult clients' self-reports (e.g., Oltmanns & Turkheimer, 2009; van der Ende, Verhulst, & Tiemeier, 2012; Vazire, 2006; Vazire & Mehl, 2008).

Second, for decades researchers and practitioners alike had a notion about the parents of youth clients, namely that when they experienced depression, their assessments are contaminated by the tendency to attend to, process, memorize, and thus report more negative behaviors about their child than other informants (i.e., depression→distortion hypothesis; Richters, 1992). Here, the implication is that depression results in parents providing inaccurate and biased reports. Stated another way, depression “clouds” a depressed parent’s judgment and decision-making capacities, and thus compromises the accuracy of the ratings they provide about their child. This is because the depression itself has little to do with what the parent rates (e.g., child’s mental health concerns) and everything to do with affective, behavioral, and/or cognitive processes that “pull” a parent’s report away from valid accounts of their child’s concerns. Thus, researchers and practitioners who accept this notion as true would either exclude that parent from the assessment process or find ways to adjust their report to “purge” it of these inherent depression-linked biases (e.g., Bauer et al., 2013; Müller et al., 2014; Youngstrom et al., 1999). Yet, one can parsimoniously link variations in parents’ levels of depression to true levels of variation in youth clients’ mental health concerns. In fact, a long line of work indicates that parental depression often portends the development of family dynamics that increase risk for mental health concerns among family members (for a review, see Goodman & Gotlieb, 1999). Further, not only is the depression→distortion hypothesis inconsistent with our understanding of the development of depression in childhood, empirical support for the hypothesis is remarkably weak (for a review,

see De Los Reyes et al., 2015). Relatedly, consider the literature on assessing pediatric bipolar disorders. Within this clinical population, parents commonly experience mood concerns themselves (e.g., Youngstrom, 2008). Thus, if the depression→distortion hypothesis were true, one would predict that this assessment literature would most likely be rife with inaccuracies in parent reports of youth mental health concerns. Yet, the opposite is true: parents provide reports of pediatric bipolar disorder symptoms that are highly predictive of the diagnostic decisions of well-trained, diagnostic consensus teams (e.g., Youngstrom et al., 2004, 2015).

These two examples illustrate that sometimes, a notion about a kind of informant's report (e.g., parent) lacking validity *falls apart* when subjected to empirical scrutiny. Given this work, you might ask: Then which informants ought to provide reports in the context of clinical services and research? We answer that question with a question. Which informants have evidence supporting their expertise as a clinically relevant information source? That is, which informants provide psychometrically sound reports of clients' behavior in the contexts in which they display mental health concerns (see also Talbott et al., 2021)? When posed this way, assessors have a considerable array of options for informants, and across clients from multiple developmental periods and presenting concerns. From early childhood into multiple periods of adulthood, assessors can often leverage clients' self-reports (e.g., Becker-Haimes et al., 2020; Hunsley & Mash, 2018; Kraemer et al., 2003). Assessors might also consider soliciting the reports of significant others in clients' lives. Specifically, decades of evidence indicate that assessments of youth benefit from including reports completed by adult authority figures, like parents and teachers, who have expertise regarding how clients behave in home and school contexts, respectively (e.g., Achenbach, 2017). Similarly, an emerging body of work now supports recommendations for assessing adult clients by soliciting reports from significant others in clients' lives, namely family (e.g., spouses), information sources in the workplace (e.g., co-workers), and non-familial caregivers in the case of elderly adults (e.g., nurses; Achenbach, 2020; Achenbach et al., 2005). Interestingly, you might assume that informants who have relevant expertise are limited to those who have a substantial length of contact with the clients about whom they provide reports. This may not be true. Indeed, recent work on assessing adolescent social anxiety finds that one can obtain psychometrically sound, incrementally valuable data from informants who base their reports on only brief interactions with the adolescent, specifically personnel trained to "stand in" as a same-age, unfamiliar peer (i.e., *peer confederate*) but who

receive no training in rating behavior (e.g., Cannon et al., 2020). In fact, more recent work supports the psychometric soundness of reports taken from unfamiliar, untrained observers (UUOs) of these same adolescents, or untrained raters who based their reports on observations of these adolescents via archived videos (Rezeppa et al., 2021).

In sum, our key goal with this section was to impress upon you a core principle that should drive decision-making regarding which informants to include in assessments. We see the informant-selection process as no different from decision-making processes for other elements of care, such as identifying interventions to address clients' needs. We suspect that it is often easy and perhaps intuitive to rely on clinical experiences and other "gut-check" heuristics when selecting informants, much like one might rely on aspects of clinical prediction when deciding on which intervention to use with a client (see also Marsh et al., 2018). To the degree that you find value in evidence-based procedures for making clinical decisions, take comfort in knowing that you have the luxury of selecting informants akin to how therapists approach selecting interventions, namely rely on the science. Rely on evidence-based standards and thresholds. After all, informants complete standardized measures, and we have well-established psychometric standards for evaluating these measures (e.g., De Los Reyes & Langer, 2018; Hunsley & Mash, 2008, 2018; Youngstrom et al., 2017). Do not divorce these standards from deciphering the value of the informants available to you. When evidentiary standards do not apply to a decision related to selecting informants, then leverage the scientific method to test any notions you might have about the veracity of a particular informant's report that you consider suspect. We suggest that you do with informants what authors for other chapters in this multi-volume resource suggest that you do with treatments. That is, accept as true notions about particular informants and the veracity of their reports only insofar as data exists to support these notions.

3.0 What Methods Do Assessors Use to Collect Informants' Reports?

Informants have much in common. For example, they live in the real world, outside of the offices and laboratories where assessors most often carry out their work. Importantly, the real world is where clients' mental health concerns and associated impairments (e.g., difficulty making friends, poor work or school performance, marital difficulties) manifest, and where practitioners aim to "move the needle" in improving mental health. Informants, who exist in the real world, are lay individuals. As such, assessors cannot assume that informants harbor expertise in how to classify and treat mental health concerns. At the same time, informants fill important

gaps in the knowledge that assessors have about their clients' mental health experiences, and this knowledge stems, in part, from their direct experiences with observing the very concerns that bring clients to seek care. In this way, the assessment process includes individuals who each carry complementary strengths and weaknesses.

The strengths and weaknesses of assessors and informants provide crucial supports to the assessment process, insofar as assessors carefully consider the “match” between the methods one uses to collect informants' reports and the ability of these informants to understand what assessors need and thus what information to provide. An exhaustive discussion of specific measures used to collect informants' reports is beyond the scope of this chapter, and there exist timely, well-written resources that provide this crucial information (e.g., Becker-Haimes et al., 2020; Hunsley & Mash, 2018). Rather, the purpose of this section is to highlight the variety of methods available to researchers and practitioners alike to collect psychometrically sound reports from informants described previously. Along these lines, in Table 1, we briefly describe commonly used multi-informant measures of mental health, personality, and adaptive functioning across the lifespan. Measures described in this table reflect well-established and psychometrically sound examples of instruments available to assessors who seek to incorporate multi-informant reports into comprehensive assessments of client mental health. As reflected in Table 1, assessors have multiple tools available to them to collect reports from informants, and here we highlight three widely used methodologies. Where possible, we point the reader to freely available instruments accessible online, to provide a direct reference to examples of measurement methods.

First, many of these reports take the form of *symptom scales*, where the informant responds to multiple, specific items (e.g., “My child fails to make eye contact when speaking to others”), each linked to a standardized response scale (e.g., 0 = never, 1 = sometimes, 2 = always) designed to assess symptoms of a specific diagnosis (e.g., ASD). The items on these surveys reference specific criteria that make up diagnostic definitions of common mental health conditions, most often those diagnoses within widely used nosological systems like the *Diagnostic and Statistical Manual of Mental Disorders* (DSM, American Psychiatric Association [APA], 2013). As a general rule, these instruments do not replace a thorough diagnostic evaluation administered by a trained professional. This is because making judgments about a client meeting criteria for a diagnosis involves multiple factors beyond the presence of symptoms, and these factors often require the judgments of trained assessors. These factors

include determining the degree to which symptoms significantly interfere with clients' functioning, and whether symptoms of a diagnosis could be explained by alternative mental or physical health conditions (e.g., Hunsley & Lee, 2014). That said, informants' reports on symptom scales often serve as screening measures that assist in determining whether additional clinical services are needed (e.g., neuropsychological evaluation) or the likelihood that clients meet criteria for a specific disorder (e.g., ASD).

Assessment of internalizing and externalizing concerns using informants' reports provide useful examples of these symptom scales. The ADHD Self-Report Scale (ASRS; Kessler et al., 2007; https://www.hcp.med.harvard.edu/ncs/ftpd/18Q_ASRS_English.pdf) assesses DSM symptoms of ADHD as they manifest with adults, and each item is an instantiation of a single symptom. Each item on the 18-item scale has its own scoring threshold for determining clinical relevance, and a summation of all item responses comprises the overall scale score. The developers of the ASRS identified a threshold on this overall score to assist in determining whether the respondent may benefit from a full diagnostic evaluation for ADHD. The ASRS includes a subset of six items that can be administered as a short version, based on research indicating that these six items best predict an ADHD diagnosis. Recently, we found that both youth self-reports and parent reports on the six-item ASRS provide psychometrically sound data regarding youth ADHD concerns (Keeley et al., 2018). Symptom scales also appear in assessments of internalizing concerns such as social anxiety. The 26-item Social Phobia and Anxiety Inventory for Children (SPAIC; Beidel et al., 1995) was created, in part, using developmentally adapted items inspired by an adult scale of the same name (SPAI; Turner et al., 1989). As with the ASRS, the SPAIC includes clinical thresholds for the total scale score. The SPAIC's developers identified this threshold based on its accuracy in detecting a social anxiety disorder diagnosis derived from a full diagnostic evaluation on an independent diagnostic interview (i.e., Anxiety Disorders Interview Schedule for Children; Silverman & Nelles, 1988).

Second, informants' survey reports are not limited to responses to symptom scales; they also contribute information on *scales about associated features* of mental health concerns. Indeed, as mentioned previously, the value of informants' reports lies in their considerable experiences with observing clients' concerns as they manifest in their social environments. These social environments also include characteristics relevant to understanding clients' concerns. These characteristics do not constitute the presence of symptoms per se, but nonetheless provide signals

that a client's clinical presentation includes characteristics known to relate to a diagnosis (i.e., *associated features*; APA, 2013), and/or pose risk to or buffer against the emergence of a diagnosis (i.e., *risk and protective factors*; Cicchetti, 1984; Luthar et al., 2000). Here, we highlight two characteristics linked to the conditions assessed by the ASRS and SPAIC/SPAI. Specifically, we have long known that the caregivers of youth experiencing ADHD often display patterns of parenting behavior that contribute to the development and maintenance of the condition, including inconsistent rule-setting and lack of praise for positive behaviors (Pelham et al., 2005). In line with this work, the Alabama Parenting Questionnaire (APQ; Frick, 1991; Frick et al., 1999) includes both parent and youth versions and assesses multiple aspects of positive and negative parenting behaviors, including parental involvement and use of harsh discipline (<https://sites01.lsu.edu/faculty/pfricklab/apq/>). Across multiple developmental periods, those experiencing social anxiety often engage in safety behaviors—subtle, maladaptive strategies such as avoiding eye contact—to minimize the distress that comes with engaging in social situations (for reviews, see Cannon et al., 2020; Piccirillo et al., 2016). Developed to assess adults, the Subtle Avoidance Frequency Examination (SAFE; Cuming et al., 2009) is a 32-item survey that assesses various safety behaviors commonly experienced by clients undergoing treatment for social anxiety (https://www.mq.edu.au/_data/assets/pdf_file/0008/596240/SAFE_English_copyright_2015.pdf). Similar to the ASRS, recent work supports use of parent and adolescent versions of the SAFE to assess safety behaviors among adolescents (Qasmieh et al., 2018), and even a modified version completed by UOs to assess these behaviors based on videotaped, archived recordings of adolescents interacting with peer confederates (Rezeppa et al., 2021).

As another example of scales about associated features, we previously mentioned that symptom scales cannot be relied upon to make judgments about diagnoses. Yet, informants nonetheless often play a prominent role in diagnostic evaluations. Using formats similar to symptom scales described previously, recent work has involved developing instruments so that assessors may collect informants' reports about clients' impairments. These include instruments such as the Work and Social Adjustment Scale (WSAS), which was originally developed to assess adults (Mundt et al., 2002), and more recently, modified youth and parent versions were developed to assess youth impairments (WSASY; De Los Reyes et al., 2019b). Both of these instruments are freely available, and the appendices of Mundt et al. and De Los Reyes et al. include the item content and scaling information for these measures.

Third, informants also provide reports in the form of responses to *interviews* administered by trained assessors (e.g., Hunsley & Mash, 2017). Assessors typically use one of a variety of interview formats, including unstructured clinical interviews which assessors can tailor to specific diagnoses (i.e., based on perceived referral needs), and standardized diagnostic interviews that include specific instructions on which diagnoses to assess and how to assess them (e.g., Rettew et al., 2009). Standardized interviews can also take a variety of formats. For instance, the Anxiety Disorders Interview Schedule (ADIS; Di Nardo et al., 1995) is a semi-structured interview designed to evaluate anxiety and mood conditions commonly experienced by adult clients. By “semi-structured” we mean that the ADIS includes specific prompts that assessors use to gauge the presence of symptoms, based on descriptions of DSM symptoms designed to be comprehensible to a layperson. Following administrations of these standardized prompts, assessors have the option of asking unstructured, follow-up questions to further probe clients’ responses to standardized prompts. Typically, these follow-up questions facilitate assessors determining factors such as but not limited to: (a) the degree to which the client understood the nature of the question and (b) whether the symptom a client endorsed manifests in more severe forms within specific contexts. For any one symptom or diagnostic judgment, the ADIS instructs assessors to base their responses on a combination of informants’ reports and clinical judgment. In contrast, within highly structured interviews such as the Diagnostic Interview Schedule for Children (DISC; Shaffer et al., 2000), informants provide responses to standardized questions about the presence of both symptoms and associated impairments, and based on these responses, built-in diagnostic algorithms determine the presence of diagnoses.

Incidentally, assessors use the interview format to collect reports from informants for purposes other than facilitating diagnostic judgments. That is, sometimes the method proves quite useful to collect reports from informants who often are incapable of providing verbal or written responses to items on traditionally formatted surveys. A key example is the Berkeley Puppet Interview (Measelle et al., 1998), which involves assessors asking young children (i.e., ages 4-8 years) to answer questions based on whether competing descriptions of behaviors as “spoken” by two puppets (i.e., one puppet worn on each hand of the assessor) is most applicable to themselves. An example of an item set might include one puppet saying to the child informant “I have lots of friends,” the other puppet saying, “I don’t have lots of friends,” and one of the two puppets asking the child “How about you?” The assessor instructs the child to respond in the way

they find most comfortable, which could include such styles as a verbal response (“The puppet on the left”), a self-description (“Nobody wants to be my friend”), and pointing to the puppet that they feel is most like them.

Fourth, with few exceptions, the measurement methods described previously focus on collecting informants’ reports based on item content linked to diagnostic systems like the DSM. Achenbach (2020) refers to these measures as based on *top-down* systems, or conceptualizations of clinical conditions, typically developed by consensus judgments offered by teams of experts and based on research and theory on mental health conditions. Achenbach contrasts these systems with *bottom-up* systems, or empirically derived syndromes consisting of multiple, statistically rare, and correlated behaviors (see also Hunsley & Lee, 2014). By “empirically derived,” we mean a collection of psychometric investigations focused on identifying syndromes that characterize maladaptive behaviors displayed among people in both the general population and clinic settings. Perhaps the most widely used bottom-up system is the Achenbach System of Empirically Based Assessments (ASEBA; Achenbach, 2017, 2020). Based on decades of research on a well-characterized set of syndromes for assessing such domains as internalizing concerns (e.g., anxious/depressed, somatic complaints) and externalizing concerns (e.g., aggressive behavior, rule-breaking), the ASEBA includes forms for assessing syndromes across development (i.e., 1 ½ to 90+ years), and as well as for developmentally appropriate informants to complete (e.g., self, parent, and teacher for youth; spouses, co-workers, and friends for adults). In sum, decades of research focused on gathering data from informants has resulted in a variety of methods for gathering multi-informant reports.

4.0 How Do Researchers and Practitioners Use Informants’ Reports?

We previously discussed the idea that informants’ reports comprise a diverse set of methods designed to balance accuracy, efficiency, and utility in assessing behaviors relevant to understanding mental health across the lifespan. Given the efficiency of these methods relative to more time-intensive methods for assessing behavior (e.g., controlled/naturalistic observations, performance-based tasks), assessors use informants’ reports for a wide variety of purposes. As with the previous section, it is outside the scope of this chapter to provide an exhaustive account of the psychometric evidence on use of informants’ reports for all available clinical and research scenarios. Here, we provide a brief overview of three areas of research and practice in which informants’ reports serve major roles.

First, in *basic research*, informants' reports factor prominently in all elements of work focused on individual differences in behavior. By "basic research," we mean studies of aspects of mental health that do not directly test clinical techniques or elements of mental health services per se, but nonetheless reveal findings that have important implications for assessing, preventing, and/or treating mental health concerns. These studies range from highly controlled laboratory research on constructs purported to explain the development of mental health concerns, to uncontrolled field studies testing the generalizability of prior findings in mental health research (e.g., in samples that are more diverse in terms of participants' background characteristics, relative to prior work). In these studies, consider the key elements of the Method section of any article you read. In basic research in clinical psychology, researchers often use informants' reports on one of the measurement methods described previously to recruit participants who all share a characteristic, such as elevated clinical symptoms of depression or specific levels of a risk or protective factor such as social skills or parenting practices. Further, informants' reports often serve as variables central to addressing research aims in basic research, including (a) predictor or independent variables, (b) criterion or dependent variables, (c) moderators or factors that modulate the relations between predictor and criterion variables, (d) mediators or factors that explain the relations between predictor and criterion variables, and (e) covariates or control variables used to determine whether relations between predictor and criterion variables result from "third variable effects" or confounding relations with covariates (see Kazdin, 2017).

Second and similar to basic research, in *applied research*, scholars implement informants' reports in all aspects of their work, only this time in studies that seek to build upon basic research findings to directly test techniques for assessing, preventing, or treating mental health concerns and/or risk and protective factors of these concerns. These studies range from randomized controlled trials testing a novel intervention, to uncontrolled studies testing the psychometric properties of a clinical instrument or moderators of treatment response when a well-established intervention is delivered in a routine clinic setting (e.g., community mental health center). In these studies, informants' reports serve as variables in all the same ways described previously for basic research. In fact, for such crucial decisions as identifying interventions supported by evidence from controlled trials, informants' reports often serve as the primary source upon which researchers make these decisions (i.e., based on their use as outcome variables in controlled trials; Weisz et al., 2005).

Third, in *practice settings*, we focus on the components of *utility* cogently described by Youngstrom (2008). Utility refers to aspects of measures above-and-beyond the psychometric properties of scores taken from them. That is, in basic research, scholars ought to ensure that measures precisely and accurately index the constructs about which scores taken from these measures ought to reflect (Nunnally & Bernstein, 1994). In practice settings and some areas of applied research, measures have to *do more* than soundly assess what they were developed to assess. In this case, informants' reports ought to facilitate decision-making regarding mental health services, and in particular improved accuracy in decision-making relative to not using a standardized measure or using an alternative measure (see also Hunsley & Mash, 2007). In this respect, Youngstrom (2008) stipulates that the utility of a measure reflects its ability to inform one or more of three aspects of service delivery. Specifically, practitioners and applied researchers use informants' reports to make *predictions* about key criteria, such as whether a client displays characteristics associated with positive responses to treatment. Informants' reports also comprise indices used to *prescribe* services, such as selecting a specific treatment to address a client's presenting concerns. Third, informants' reports are used to make decisions regarding therapeutic *processes* once a client is already undergoing treatment, such as whether a report indicates that the client is at risk for dropping out of treatment or should be discharged from treatment.

Importantly, relatively few measures, informants' reports or otherwise, demonstrate evidence of utility (see also De Los Reyes & Langer, 2018; Hunsley & Mash, 2007). One of the few examples involves tests of the utility of the Outcome Questionnaire-45 (OQ-45), a weekly self-report measure used to index adult clients' psychosocial functioning during treatment (Lambert, 2007). A series of sophisticated, controlled experiments leveraged the OQ-45 to test whether the week-to-week administration of the instrument facilitates making better clinical decisions related to clients' functioning (e.g., Lambert et al., 2003). In these studies, therapists were randomly assigned to one of two groups: (a) access to weekly OQ-45 data for their clients, coupled with feedback as to whether each client deteriorated, remained stable, or improved in functioning from week to week; or (b) access to assessments "as usual," or weekly OQ-45 data that did not include feedback about treatment progress. A prior quantitative review of these controlled trials indicates that, on average, OQ-45 data coupled with the feedback system reduces rates of client deterioration relative to assessment as usual (Shimokawa et al., 2010). This

research provides a model for how to test the utility of informants' reports and potentially, how scores taken from these reports may facilitate decision-making in practice settings. For instance, using OQ-45 data coupled with the feedback system, practitioners can detect clients whose functioning is deteriorating and subsequently modify the treatment approach to address these deteriorations. Overall, researchers and practitioners leverage informants' reports for numerous purposes and throughout various stages of research (i.e., basic and applied), and clinical work (e.g., classification and diagnosis, treatment planning, monitoring treatment response).

5.0 From How Many Informants Do Assessors Collect Reports?

By now, you have probably surmised that when it comes to an assessor deciding on the number of informants from whom they should collect reports, a safe answer is, "More than one." In fact, "best practices" in clinical assessments of youth mental health have long incorporated a multi-informant approach to assessment (e.g., De Los Reyes, 2011; Hunsley & Mash, 2007). Yet, beyond collecting reports from more than one informant, no guidelines exist in *any* assessment literature for making even rudimentary decisions with regard to multi-informant approaches to assessment (see also Beidas et al., 2015; De Los Reyes et al., 2019c). For example, we lack guidelines for upper limits on the number of informants from whom to solicit reports in research and/or practice settings. Consequently, we know little as to whether multi-informant assessments ought to include two, three, four, or as many as five informants. It is reasonable to assume that across literatures, the incremental value of adding an informant decreases as the number of informants increases. That said, the answer to this question regarding the specific number of informants likely varies across settings. Further, the feasibility of collecting multi-informant data might vary depending on whether the assessments are conducted for research versus clinical purposes. As such, developing guidelines for collecting, using, and interpreting reports from multiple informants comprises a crucial area of study in clinical assessment. Interestingly, multi-informant scenarios in assessment literatures outside of clinical psychology (e.g., 360° feedback approaches to performance evaluations in organizational behavior research) leverage paradigms that frequently involve five or more informants (e.g., reports from supervisors, subordinates, and colleagues at similar ranks or job titles; Brett & Atwater, 2001).

Research on *incremental validity* reveals challenges with developing a set standard on upper limits for the number of informants' reports in an assessment. Incremental validity focuses on whether scores taken from measures explain variance in clinically relevant criterion variables,

over-and-above variance explained by alternative measures (Borsboom, Mellenbergh, & van Heerden, 2004; Hunsley & Meyer, 2003). This focus reflects the fact that all assessors have finite time and material resources available to carry out assessments (e.g., costs of instruments, costs to train personnel; Hunsley & Mash, 2007). Thus, the degree to which an assessor includes a specific measure in an assessment hinges, in part, on whether scores taken from that measure reveal meaningful aspects of client functioning that cannot be obtained from scores taken from other measures included in the assessment (see also Blais et al., 2001). In these respects, relevant to the construction of any multi-informant assessment is the question of what expertise an informant brings to an assessment (e.g., their opportunities for observing the client in a context relevant to their care), and whether that expertise complements the expertise of the other informants involved (see also Dawes, 1999; De Los Reyes et al., 2015).

The incorporation of incremental validity research in the informant selection process reveals a seemingly straightforward process, namely to set a “bar” of incremental value and complementary expertise that all informants should pass. Yet, two key elements of incremental validity research present important nuances to this decision-making process. First, an important aspect of incremental validity research is the observation that conclusions about a measure’s incremental validity change depending on the criterion variable used (Haynes & Lench, 2003). For instance, parent reports might display incremental validity in predicting youth diagnostic status but not youth academic achievement, above-and-beyond teacher reports. Under these circumstances, how does one decide whether parent reports demonstrate sufficient (or insufficient) incremental value to include in an assessment? The answer to this question depends on the purpose of the assessment and the link between that purpose and the nature of the incremental validity evidence.

Second, we previously discussed the need to determine potential overlap between information gleaned from each informant’s report used in a multi-informant assessment. In line with this consideration, determining each informant’s incremental value requires close attention not only to the overlap among informants’ reports, but also the criterion variables used to test incremental validity. Specifically, it is important to avoid overlap between methodology and/or information sources used to capture the criterion variable (e.g., parent reports on child’s symptoms *and* criterion variable measure). When this *criterion contamination* occurs, researchers can inadvertently “stack the deck” in favor of or against a particular informant who shares

method variance with the criterion variable (Garb, 2003). A common scenario in which criterion contamination occurs is with use of diagnostic status as a criterion variable. Here, depending on the diagnostic procedures, some but not all informants may be involved in the process of compiling information that results in the diagnosis (see also De Los Reyes et al., 2015). Thus, parent reports may display incremental validity in predicting diagnostic status, over-and-above teacher reports, in part, because the study design involved relying only on parents to provide reports to inform diagnostic judgments (e.g., Bossuyt et al., 2003; Whiting et al., 2011).

Taken together, work on the incremental validity of informants' reports may assist in making determinations as to which and how many informants to include in an assessment. Yet, the design of incremental validity studies and in particular the selection of criterion variables weighs heavily in the findings of this work, and thus the use of this evidence to guide decision-making. The following sections reveal key lines of theoretical and empirical work that reveal additional challenges inherent in this decision-making process, as well as a path toward potential solutions, recommendations, and future directions.

6.0 What Happens When More Than One Informant Reports about the Same Client?

Our review thus far of the literature on informants' subjective reports of mental health reveals several themes, including: (a) informants vary in their experiences or observations of clients within specific contexts (e.g., home vs. school/work vs. peer interactions); (b) clients vary in the contexts in which they display mental health concerns; and thus (c) a comprehensive picture of clients' functioning within and across relevant contexts requires collecting reports from, at minimum, more than one informant. Yet, these themes are insufficient to justify taking a multi-informant approach to assessment. In line with our discussion of incremental validity, a crucial element of gathering informants' reports involves understanding the value each informant brings to the assessment. That is, is there evidence to support the notion that informants provide non-redundant reports, such that one cannot accurately estimate what one informant might report, based on data from another informant's report?

Perhaps the strongest evidence base supporting the lack of redundancy in multiple informants' reports lies in research estimating levels of correspondence (in most cases estimates based on the Pearson r metric) between informants' reports of the same target individual (e.g., parent and teacher reports of the same child's ADHD symptoms). This literature traces back to the 1950s (Lapouse & Monk, 1958) and continues to the present day (De Los Reyes & Makol, in

press). In fact, meta-analyses have estimated cross-informant correspondence levels across assessments of youth and adults, as well as across assessments of mental health domains and their associated features. For example, Achenbach and colleagues (1987) reported an overall Pearson r estimate of .28 for 119 studies using cross-informant reports of youth mental health. An independent meta-analysis of 341 studies on correspondence among informants' reports of youth mental health published between 1989 and 2014 revealed an *identical* overall estimate ($r = .28$; De Los Reyes et al., 2015). Recent discussions point to a key limitation of psychological research generally, namely that it reflects psychological phenomena as displayed by individuals with a specific, nearly homogenous set of characteristics (e.g., Henrich et al., 2010). In line with these discussions, you would be right to ask how representative cross-informant correspondence effects might be. That is, to what degree does this low-to-moderate level of correspondence apply to multi-informant assessments as conducted within populations of youth receiving mental health care globally? Interestingly, a recent meta-analysis revealed that multi-informant correspondence studies have been conducted in over 30 countries, spanning over six continents (De Los Reyes et al., 2019a). In this meta-analysis, the 95% confidence interval for correspondence as estimated using the Pearson r metric revealed a surprisingly tight interval (.24, .31) across these cross-cultural studies. The replicability of these estimates—across hundreds of studies conducted globally—rivals the replicability of placebo effects (cf. Ashar et al., 2017).

Beyond cross-cultural relevance, you might also be curious as to whether cross-informant correspondence estimates manifest in the same or similar ways across assessment conditions. Indeed, in other chapters in this multi-volume resource, you will read a great deal about how effects observed for interventions in highly controlled conditions (e.g., university laboratories) often attenuate considerably when tested in less controlled conditions (e.g., community clinics; Hunsley & Lee, 2014). Interestingly, not only do robust, low-to-moderate levels of correspondence manifest cross-culturally, these correspondence levels appear across assessment settings (e.g., laboratory clinic vs. community clinic) and regardless of whether the assessment involved clinic or community samples (for a review, see De Los Reyes et al., 2015). Further, a series of additional meta-analyses support the replicability of these low-to-moderate levels of correspondence. Specifically, these levels of correspondence manifest across meta-analyses of assessments of adult mental health (Achenbach et al., 2005), youth autism (Stratis & Lecavalier, 2015), parenting (Hou et al., 2019; Korelitz & Garber, 2016), youth social competence (Renk & Phares,

2004), and maltreated youth (Romano et al., 2018), among numerous other areas. Collectively, the meta-analyses described previously provide estimates of correspondence based on data from well over 500 studies, and reveal a remarkably strong, robust estimate of low-to-moderate levels of cross-informant correspondence in reports collected within clinical assessments. Thus, the question becomes not what one should do *if* these *informant discrepancies* arise but rather what one should do *when* they arise.

7.0 Using Multiple Informants' Reports When Interpreting Research Findings

In the previous section, we described research that consistently reveals low-to-moderate levels of cross-informant correspondence in clinical assessments. These correspondence levels manifest not only across hundreds of studies, but also studies across a diverse array of assessment settings, and in assessments administered globally. Yet, do these correspondence estimates mean that the actual *findings* of research change, depending on the informant? Indeed, so much of research involves making discrete judgments or conclusions based on data. Consider the following questions:

- *What is the 12-month prevalence of anxiety disorders in the United States?*
- *Are certain cognitions or avoidance behaviors risk factors for anxiety disorders?*
- *Is cognitive behavioral treatment efficacious in reducing symptoms of anxiety disorders?*

Each of these represents an important research question that, depending on the answer provided by a study addressing the question, may point to one of various directions for future research and practice. Stated another way, all studies potentially shape the focus of future studies and perhaps eventually, the delivery of mental health services in practice settings. In this respect, the source of data pointing to a research finding has the potential to influence entire lines of research for years to come. It is here where the implications of informant discrepancies truly come into focus.

The three questions highlighted in the previous paragraph reflect distinct points in a line of clinical research: (a) rates of mental health concerns, (b) associated features of mental health concerns, and (c) effectiveness of mental health interventions. In many lines of research, each of these points builds on the other, such that “a studies” identify needs for intervention, “b studies” inform the development of interventions, and “c studies” test these interventions. Importantly, research indicates that in each of these kinds of studies, changing the informants used to address study aims often changes the conclusions drawn from these studies (for reviews, see De Los Reyes & Kazdin, 2005, 2008).

For example, in an early study, the prevalence of conduct and oppositional defiant disorders in a representative, general population sample ranged from 1.6% to 10.2%, depending on whether rates reflected a specific informant's report (i.e., parent or teacher) or reflected data from an algorithm that combined the reports (i.e., AND/OR rule; Offord et al., 1996). A more recent study focusing on oppositional defiant disorder reflected wide ranges, albeit with lower absolute values (i.e., 0.2% to 2.6%; Munkvold et al., 2009). These ranges of prevalence by informant also appear in studies of clinic samples. In a study using parent and teacher ratings (and their combination), prevalence of conduct disorder ranged from 9.7% to 23% and prevalence of anxiety and depressive disorders ranged from 10.3% to 36.2% (MacLeod et al., 1999). One observes even wider prevalence estimates when assessing comorbidity: based on parent, youth, or teacher ratings (or their combination) of a sample of outpatient youth, comorbidity rates ranged from 5.4% to 74.1% (Youngstrom et al., 2003). In each of these cases, changing the informant changes the conclusions one draws as to how many individuals display mental health concerns and thus, the overall need for mental health services linked to these concerns.

Studies of associated features of mental health concerns reveal similarly inconsistent findings as a function of the informant. For instance, are certain thought processes associated with a diagnosis of childhood depression? The answer to this question largely depends on which informant (i.e., parent vs. child) provides reports about these thought processes, as well as to determine whether the child is depressed (Kazdin, 1989). Further, while conduct disorder based on parent reports relates to parent depression and family dysfunction, conduct disorder based on teacher reports relates to child gender and family income. However, these relations are mutually exclusive of each other, such that characteristics relate to either parent- or teacher-identified conduct disorder but not both (Offord et al., 1996). The general finding that one identifies associated features of mental health concerns based, in part, on which informants provide data on associated features and mental health domains is also reflected in more recent work (e.g., Dirks et al., 2011; Drabick et al., 2008). As with prevalence rates, changing the informant changes the conclusions one draws as to the associated features of mental health concerns. This is a key issue, as research on these features often informs the development of intervention techniques (e.g., parent training for youth conduct problems; Kazdin & Rotella, 2009).

Taken together, prior work indicates the potential to come to widely different conclusions as to *who* experiences mental health concerns as well as *what factors* are either associated with or

perhaps even contributing to the maintenance of mental health concerns, depending on the informants used in these studies. Importantly, we previously mentioned that these same informants also provide reports used to gauge the efficacy of interventions (e.g., Weisz et al., 2005). Thus, not surprisingly, meta-analytic reviews of intervention outcomes consistently reveal that estimates of intervention effects vary as a function of informant. For instance, in the literature on youth treatments, diverging findings as to intervention effects commonly manifest as a function of informant (e.g., Casey & Berman, 1985; De Los Reyes & Kazdin, 2009; Weisz, Weiss, Alicke, & Klotz, 1987; Weisz, McCarty, & Valeri, 2006). In fact, across these reviews, estimates of intervention effects range from “small” to “large” when guidelines to interpret these effects are based on effect size conventions in the social sciences (e.g., Cohen’s d ranging from 0.1’s to 1.0+; Cohen, 1988). These variations by informant occur with effect size estimates for adult interventions as well (e.g., Cuijpers et al., 2010, 2011, 2014; Lambert et al., 1986; Ogles et al., 1990; Renner et al., 2014). Importantly, many of the same intervention studies in these meta-analytic reviews constitute the evidence used to identify interventions that have displayed sufficient efficacy to be recommended for use in routine clinic settings (i.e., evidence-based interventions; see De Los Reyes, 2011; De Los Reyes & Kazdin, 2006, 2008, 2009).

8.0 Using Multiple Informants’ Reports When Making Clinical Decisions

Similar to researchers, clinicians leverage informants’ reports to make important decisions regarding care. Across several different points of care, clinicians often encounter low correspondence among informants’ reports. For example, when determining problems to target in treatment, parents and children *fail to agree* on a single target problem 63% of the time and fail to agree on even the general category of target problems 36% of the time (e.g., aggressive behavior; anxiety/depression; Yeh & Weisz, 2001). Considering therapists’ impressions of treatment targets along with those of parents and children reveals even lower levels of correspondence, with over 76% of parent–child–therapist triads failing to agree on a single target problem, and over 44% failing to agree on a general category of problems (Hawley & Weisz, 2003).

Importantly, when confronted with discrepant informant reports, clinicians are still faced with using these reports to make important decisions regarding care (e.g., treatment planning). Research indicates that clinicians tend to make clinical decisions that emphasize one informant’s reports over others’ reports (for a review, see De Los Reyes et al., 2015). That is, rather than integrate data gained from all informants used in an assessment, they tend to “take sides” and rely

only on one informant's report. Recently, we examined this phenomenon using an experimental design in which we exposed clinicians to vignettes that varied as to which informants' reports indicated greater improvement following treatment (e.g., parent vs. child vs. teacher; Marsh et al., 2020). We found that clinicians' weighting of informants' reports depends on the domain assessed and the informant pair examined. For example, clinicians tend to make judgements that rely more on the child report than parent report for decisions about children's internalizing problems, and more on the parent report for decisions about externalizing problems. In this respect, clinicians appear to make judgments in line with the assumption that informants vary as to whether they have sufficient insight about the target of treatment and consequently disregard the informant who they believe lacks insight.

The idea that clinicians often fail to integrate informants' reports when making clinical decisions is a problematic one, especially considering that there is no evidence indicating that any one informant provides "better" or more accurate information about any mental health domain, relative to other informants (see also De Los Reyes, Kundery, & Wang, 2011). We would argue that these effects likely manifest because clinicians lack any guidelines for using and interpreting the outcomes of multi-informant assessments. Incidentally, researchers are not immune to these issues and they too lack guidance on using and interpreting multiple informants' reports. This is likely why many researchers leverage approaches to analyzing or modeling data that appear inconsistent with the latest research on what low correspondence among informants' reports likely reflects (for a review, see De Los Reyes et al., 2019a). In the following section, we provide an overview of the latest efforts to build conceptual models of multi-informant assessments, along with research informed by these models.

9.0 Conceptualizing Patterns in Multiple Informants' Reports:

The Operations Triad Model

Advancing use and interpretation of multiple informants' reports begins with hypothesis testing. That is, to transform clinical decision-making when using informants' reports from an idiosyncratic to an empirically based process requires building an evidence base. The Operations Triad Model (OTM; De Los Reyes et al., 2013) offers a framework for building this evidence base (see Figure 1). Specifically, the OTM provides a guide for researchers and clinicians on how to form a priori hypotheses about the degree to which informants' reports will converge and/or diverge. The

OTM is agnostic as to whether patterns among informants' reports have utility, and can be used to test questions such as:

- *Does disagreement between self-reports and coworker reports of interpersonal functioning reflect meaningful variations in functioning across contexts?*
- *Does agreement between parent and teacher reports of youth ADHD symptoms indicate that youth are more impaired, relative to instances in which reports disagree?*

Given that assessment conditions vary widely by the domain(s) assessed, informants used, and settings, one can assume that informant discrepancies will reflect meaningful information in some but not all cases. Below, we review the OTM to illustrate how it can be used to evaluate the clinical utility of multi-informant assessments and enhance interpretation of patterns among informants' reports.

The OTM includes three measurement conditions that guide hypothesis testing about the meaning underlying patterns among multi-informant reports. First, *Converging Operations* and *Diverging Operations* reflect measurement conditions in which patterns of convergence and divergence among informants' reports reflect meaningful information. *Converging Operations* reflects a set of measurement conditions for interpreting patterns of consistent reports in accurately reflecting the same conclusion or yielding useful clinical information. For example, a parent and teacher may both report elevated youth anxiety symptoms when a child displays anxiety at both home and school or has a more severe symptom presentation. In contrast, *Diverging Operations* reflects a set of measurement conditions for interpreting patterns of inconsistent information sources based on hypotheses about variations in the behavior being assessed. For example, a teacher may provide higher anxiety ratings than a parent given that the child displays anxiety when interacting with peers in the classroom but does not display anxiety at home. In these examples of *Converging* and *Diverging Operations*, useful information is gleaned when interpreting patterns of informants' reports, with implications for clinical decision-making tasks such as diagnosis and treatment planning. Importantly, when evidence for *Diverging Operations* is found, researchers and practitioners would want to avoid using techniques that rely exclusively on maximizing "common variance" among informants' reports and minimizing or ignoring "unique variance" (i.e., discrepancies between reports). One example of this approach is the AND/OR rules described previously. Rather, assessors ought to leverage

techniques that capitalize on information derived from these differences to enhance the assessment information and task at hand. Below, we will describe several of these techniques.

Sometimes, discrepancies among informants' reports do not reflect Diverging Operations conditions, but rather conditions explained by methodological features of the assessment tools (e.g., report forms) or informants (e.g., malingering) being used, what the OTM terms as *Compensating Operations*. Historically, the prevailing assumption about informant discrepancies focused on such measurement conditions, with most interpretations focused on informant bias or measurement error as explanations (De Los Reyes, 2011). As with Converging and Diverging Operations, identifying conditions that reflect Compensating Operations requires evidence. There are several aspects of the measures being used that researchers and practitioners can consider when identifying conditions that reflect Compensating Operations. For example, informant discrepancies may arise when one uses measures with unique item content, reliability, scaling, or scoring, even if the measures purportedly assesses the same construct. For example, when assessing youth depression, a practitioner may administer completely distinct symptom measures to children (e.g., Berkeley Puppet Interview; Measelle et al., 1998) and their parents (e.g., Children's Depression Inventory [CDI]; Kovacs, 1992). Key to Compensating Operations is that differences among informants' reports do not reflect meaningful information and thus informant discrepancies themselves cannot enhance the utility of information gathered in an assessment. In fact, evidence supporting Compensating Operations also supports use of techniques that maximally focus on common variance and minimize or ignore unique variance, such as AND/OR rules or structural equations modeling (e.g., see also Holmbeck et al., 2002)

There is a large and growing evidence base supporting Converging and Diverging Operations hypotheses. We highlight particularly exciting work in this area, and a full review is beyond the scope of this chapter (for reviews, see De Los Reyes et al., 2019a, 2020). By construction, OTM studies take the comprehensive assessment approach we described at the opening of this chapter. Within these studies, researchers often use well-established analytic techniques (e.g., latent class analysis; LCA; see McCutcheon, 1987) to test for the presence of two or more patterns of multi-informant reports. These studies tend to identify multiple reporting patterns within the overall sample, indicating that, although overall correspondence levels between informants' reports tend to fall in the low-to-moderate range, this does not signal that all informants' reports disagree in the same way, or even that two informants' reports never agree

(see also De Los Reyes & Ohannessian, 2016). In fact, in large samples one learns that informants' reports agree and disagree in many ways. Specifically, researchers often observe patterns characterized by convergence or divergence between reports including: (a) classes characterized by convergence in high or low levels of symptom reports by informants, and (b) classes in which one informant reports higher levels of symptom reports than the other informant (and vice versa; Makol et al., 2021).

After identifying reporting patterns, OTM-informed research involves testing hypotheses about the utility of informants' reporting patterns by evaluating how these patterns relate to independent criterion variables (e.g., diagnosis, impairment, treatment outcomes). An area of research with perhaps the most support finds that informants' reports relate to cross-contextual variations in behavior. De Los Reyes and colleagues (2009) found that parent-teacher reports of young children's disruptive behaviors are associated with children's observed disruptive behavior in the home and school contexts. More specifically, children were likely to display elevated disruptive behaviors in the specific contexts in which the informants reported elevated disruptive behavior and were unlikely to when informants did not (e.g., child displays home-specific disruptive behavior when only the parent reports elevated disruptive behaviors). Research on multi-informant assessment of social functioning finds that teacher-caregiver reports of youth aggression and social withdrawal relate to social events encountered by youth across contexts (Hartley, Zakriski, & Wright, 2011), teacher-peer reports of youth social skills each provide incrementally valid data for predicting context-specific social functioning (Kwon, Kim, & Sheridan, 2012), and adolescent-unfamiliar peer reports (but not parent reports) of adolescent social anxiety relate to adolescents' perceived arousal in social interactions with unfamiliar peers (Deros et al., 2018). Although less is known about informant discrepancies when assessing adult mental health, research suggests the association between informants' reporting patterns and cross-contextual variations in behaviors extends through adulthood. For example, convergence in clinician and adult clients' reports of elevated social anxiety symptoms is associated with social skills deficits across social interaction contexts (De Los Reyes, Bunnell, & Beidel, 2013). Overall, research using the comprehensive assessment approach demonstrates that clients vary in the contexts in which they display behaviors and informants across these contexts have unique opportunities to observe clients' behaviors.

Emerging work using the comprehensive assessment approach supports the notion that patterns among informants' reports yield information beyond contextual variations in behavior. For example, patterns among parent-adolescent reports of depression and suicidal ideation relate to youth mental health service use history (Jones et al., 2019; Makol & Polo, 2018). Some research suggests that divergence among informants' reports (i.e., one informant reporting elevated clinical concerns while the other does not) is associated with unique risk. Specifically, Lippold and colleagues (2013, 2014) found that higher parent than adolescent report of parental monitoring of adolescent activities is associated with increased adolescent risk for developing substance use problems over time. Patterns among informants' reports at the start of treatment may also predict treatment characteristics. Becker-Haimes et al., (2018) found that youth self-reporting lower levels of anxiety symptoms than their parents at the start of treatment are less likely to make progress in therapy. In addition, Makol et al. (2019) found that patterns among parent-adolescent reports of internalizing problems at the start of psychiatric inpatient treatment are associated with diagnosis, treatment engagement, and length of stay on the unit. Overall, OTM research supports that examining patterns among informants' reports yields information of use in many assessment tasks and across populations and domains.

10.0 Approaches to Integrating Multiple Informants' Reports

Even when the evidence indicates that multi-informant reporting patterns reflect meaningful clinical information, researchers and practitioners still face a complex task: how to integrate the reports to inform sound clinical decision-making. As we describe below, multiple approaches exist for integrating multi-informant data, although we wish to highlight two key issues at the outset of describing these approaches. First, the state of the science on evidence-based approaches to integrating multiple informants' reports currently applies to interpreting these data at the sample level, an issue to which we return later on in the chapter (see 11.0 Future Research Directions; 12.0 Clinical Applications and Recommendations). Second, something important to understand about all integrative approaches is that they draw on unique assumptions about the meaning underlying informants' reports. Unfortunately, we have a nascent evidence base to guide the development of empirically derived and clinically useful integrative approaches. It is likely that multiple integrative approaches will enhance use of multi-informant reports, and that the approach taken will vary depending on the rationale for using multi-informant data (e.g., interpreting research findings, making clinical decisions). No one researcher or practitioner can avoid the task of integrating multi-

informant data. That is, the very reason why we take a multi-informant approach to assessment is that for no assessment setting can one point to a “gold standard” or single data point that can guide all decision-making (see also De Los Reyes, Augenstein, & Aldao, 2017). Thus, it is incumbent on users of integrative strategies to scrutinize the assumptions and evidence base inherent in leveraging the strategy.

Several common approaches for “reconciling” discrepant informant reports undermine the very purpose of collecting these reports. We describe the most prominent of these approaches here to aid the reader in being a critical consumer of integrative approaches, including those that seem intuitive but lack empirical support. First, to reconcile informant discrepancies, it is common to choose a single “optimal” index, an approach commonly taken in randomized controlled trials of treatments (i.e., primary outcome measure; De Los Reyes et al., 2011). Importantly, sometimes independent evaluators’ ratings serve as this singular index, and prior work indicates that these evaluators are more likely to make clinical decisions (e.g., form diagnoses) that are consistent with parent reports when they encounter discrepancies between their reports and those of other informants (e.g., adolescents; Brown-Jacobsen, Wallace, & Whiteside, 2011; Hawley & Weisz, 2003). Because the evidence does not support use of single, optimal informants in youth mental health assessments, taking this approach results in using subjective decision-making processes that effectively discard multi-informant data. Alternatively, one may decide to analyze informants’ reports separately to arrive at a conclusion (e.g., conducting separate statistical tests with parent, teacher, and youth reports of anxiety). Although this approach aims to utilize information from all informants involved in an assessment, one still faces the challenge of understanding how to come to a sound conclusion when encountering informant discrepancies (e.g., when parent reports of anxiety relate to an outcome, but teacher and youth reports do not). Further, analyzing each informant’s report separately from the reports of other informants can lead to increases in both Type I errors (i.e., increased likelihood of false positive results due to not adjusting for multiple tests) and Type II errors (i.e., increased likelihood of false negative results after adjusting for multiple tests; Kraemer et al., 2003). Alternatively, one may opt to use combinational rules or algorithms to arrive at a clinical decision (e.g., when using parent and teacher reports to arrive at an ADHD diagnosis). For example, combinational rules or algorithms such as the “*OR*” rule (i.e., symptom is considered “present” if any informant endorses it) and “*AND*” rule (i.e., symptom is considered “present” when both informants endorse it; Bird et al. 1992; Jensen et al. 1999; Piacentini et al. 1992). However, there is no evidence to suggest that combinational rules and algorithms lead to improved clinical decision-making (De Los Reyes et al., 2015). As mentioned previously, these

approaches remove information about which informant(s) reported elevated concerns, and thus they effectively remove context-specific information and the ability to detect context-specific associated features (e.g., Offord et al., 1996).

Finally, one may opt to integrate informants' reports using a composite variable approach (e.g., taking an average of parent, teacher, and youth reports of youth anxiety). However, this approach assumes that each informant's report reflects "error" around a "true score" representation of the construct being assessed, thus treating differences among reports as error (Borsboom, 2005). As noted previously, there is a large evidence base supporting that differences among informants' reports reflect meaningful information. Thus, it is unsurprising that the composite score reduces predictive power and the unique contributions of each informant (Makol et al., 2020). When considering use of these integrative approaches—optimal informants, separate statistical analyses, combinational rules and algorithms, and composite score approaches—we encourage the reader to ask what is gained, and perhaps more importantly, *what is lost* when using one of these approaches. That is, does the approach enhance the information provided by informants or minimize them? Whenever possible, we encourage the reader to examine the empirical support for the integrative approach they select, and particularly by drawing on research demonstrating that the approach results in incremental validity and the ability to explain variance in clinically relevant criterion variables (e.g., diagnosis, treatment outcome).

So what works when integrating informants' reports? Informants' reports provide complex information, and integrating them is thus a complex task. Given that this is an emerging research area, we highlight several promising approaches. As previously mentioned, LCA represents a person-centered approach to identifying reporting patterns within an overall sample. This approach has been applied to informants' reports across diverse problem types, developmental periods, and informants (De Los Reyes et al. 2009, 2016a; Lerner et al. 2017; Lippold et al. 2013, 2014; Sulik et al., 2017; Makol et al., 2019). For example, when using LCA to characterize parent and teacher reports of youth autism symptoms, Lerner and colleagues (2017) found that convergence in informants' reports of elevated autism symptoms was associated with psychotropic medication use, special education service enrollment, and an autism diagnosis using a well-established assessment tool. A useful approach for understanding the relation of informant discrepancies to outcomes is the polynomial regression approach (Edwards, 2002). This approach allows one to control for the main effects of each informant's reports on outcomes while also isolating the unique effects of informant discrepancies (Laird & De Los Reyes, 2013). Research leveraging this approach has demonstrated that informant discrepancies predict

unique variance in outcomes above-and-beyond the variance explained by individual informants' reports (Becker-Haimes et al., 2018; Laird & De Los Reyes, 2013; Nelemans et al., 2016; Xu, Boyd, Butler, Moore, & Benton, 2017). For example, Becker-Haimes and colleagues (2018) found that informant discrepancies characterized by lower youth relative to parent report of youth anxiety at the start of treatment portends poorer youth treatment outcomes.

Finally, Kraemer and colleagues' (2003) "Satellite Model" can be used both to identify the appropriate informants to select in clinical assessments and efficiently integrate them. Kraemer and colleagues (2003) argue that effective use of informants' reports begins with careful selection of the informants included in the assessments. Specifically, they recommend using their "mix-and-match criterion" to select informants who systematically vary in the context (e.g., home, school) and perspective (e.g., self, other) from which they provide reports. This approach rests on the assumption that one should select informants who can be expected to agree, so as to enhance the clinical utility of an assessment. The Satellite Model leverages principal components analysis (PCA) to parsimoniously integrate informants' reports. This approach yields a *Trait* component, which captures concerns that manifest across informants' contexts and perspectives. Makol et al. (2020) recently applied the Satellite Model to multi-informant assessments of adolescent social anxiety. In this study, informants were selected for their unique contexts and perspectives using the "mix-and-match" criterion: (1) parents and peer confederates were selected given that they each observe adolescent behavior from an other-perspective but from unique contexts (i.e., home vs. peer environments) and (2) adolescents were selected given that they observe their behavior from a self-perspective and across both home and peer contexts. When integrating these informants' reports of social anxiety using PCA, Makol and colleagues (2020) found that the Trait score optimized prediction of independent criterion variables. Specifically, the Trait score outperformed individual informants' reports and a composite score of informants' reports in predicting clinical variables important for characterizing adolescent social anxiety (i.e., observed adolescent anxiety in social interaction tasks, adolescent referral status), representing medium-to-large effects. We encourage readers to leverage approaches such as LCA, polynomial regression, and the Satellite Model to integrate informants' reports. We surmise that using these approaches, and other research-supported integrative strategies, will optimize informants' reports while also avoiding the loss of valuable information.

11.0 Future Research Directions*

The work we reviewed throughout this chapter opens doors to exciting directions in research, of which we will highlight three. First, we noted that informant discrepancies manifest in mental health assessments with clients throughout the lifespan. However, the evidence that supports links between informant discrepancies and contextual variations in mental health concerns largely comes from research on assessments of youth mental health (for an exception, see De Los Reyes et al., 2013b). Thus, future work ought to apply the OTM and laboratory-based paradigms informed by this framework to understand the degree to which variations in informant discrepancies signal contextual variations in mental health concerns among adults.

Second, even in youth assessments—where substantial evidence suggests links between cross-informant correspondence and contextual variations in mental health concerns—evidence documenting these links largely comes from cross-sectional work. Do changes in the levels of informant discrepancies *over time* signal shifts in the specific contexts where clients display mental health concerns? We cited some research to indicate that levels of informant discrepancies predict important clinical outcomes (e.g., treatment response; Becker-Haimes et al., 2018; Makol et al., 2019), providing circumstantial evidence to indicate that these discrepancies yield information about future functioning. Extending this work to the longitudinal study of informants' reports may greatly inform our understanding of how context factors into the development, maintenance, and treatment of mental health concerns. In line with this, research supports use of some of the analytic techniques cited in this chapter for understanding longitudinal processes. In particular, recent advancements in use of polynomial regression techniques now allow for studying the variables used to index informant discrepancies (i.e., interaction terms) as both predictor and criterion variables, allowing for studying longitudinal changes in such discrepancies (De Los Reyes, Ohannessian, & Laird, 2016b; & LaFleur, 2016). Similarly, person-centered models such as LCA include variants that allow for detecting changes in class assignment over time (i.e., latent transition analysis; see Lippold et al., 2014). We recommend that future research leveraging these and other approaches focus on understanding the development of informant discrepancies and their links to longitudinal changes in key aspects of mental health, such as the specific contexts in which individuals experience mental health concerns.

Third, all of the work we reviewed in this chapter focuses on samples of individuals. When studies identify links between informant discrepancies and information germane to mental health, this work assists in using and interpreting multi-informant data at the sample level. Yet, how might theoretical and measurement models about informant discrepancies operate with individual clients? For instance, how might studies that leverage data from the Kraemer and colleagues (2003) Satellite Model inform interpretations of multi-informant data collected about individual clients? To address questions about individual clients, a crucial next step may involve leveraging applied measurement techniques to multi-informant data. Specifically, researchers who develop symptom scales often recruit large, representative community samples and/or samples of clients to identify *normative scores*, or scores used to distinguish individuals with clinically elevated symptoms from individuals whose symptoms fall in a “typical” range of healthy functioning (see Kazdin, 2017). Using these methods of normative scoring, an interesting direction for future research might involve applying the Satellite Model to large samples of multi-informant data. Based on prior work that indicates scores derived from the Satellite Model provide incrementally valuable information in relation to independent criterion variables (e.g., Trait score), follow-up work on the informants used in these studies could “scale up” the sample size to identify normative scores to be applied to individual cases. Examples of how to carry out these procedures already exist for creating normative scores for instruments administered to individual informants (e.g., Achenbach, 2020; Kessler et al., 2007). We recommend that researchers leverage these same procedures to arrive at normative scores that are derived from integrated multi-informant data, and use approaches that have undergone rigorous validation testing. This work could transform use of multi-informant data in routine clinic settings, where practitioners commonly use multi-informant assessments with individual clients. How might these processes manifest in routine settings? We address this question in the next section.

12.0 Clinical Applications and Recommendations*

As discussed throughout this chapter, informant reports represent an important component of clinical assessment. Clients lead complex lives and we leverage informant reports to capture at least some of this complexity. Research on the validity of scores taken from informant reports and strategies for integrating them has implications for clinical practice. We highlight these implications with two recommendations. First, soliciting informants for use in clinical assessments, we recommend that clinicians thoughtfully consider *whose report* is important to

leverage and *what data* each informant provides. For example, when a clinician is making diagnostic decisions for psychiatric disorders requiring assessment of multiple contexts (e.g., ADHD, ASD), they may collect information from a client as well as their parent and teacher (De Los Reyes et al., 2015). As another example, when assessing personality disorders, a clinician may consider collecting information from their client as well as their client's coworker to understand the intrapersonal and interpersonal aspects of the client's functioning (see Carlson, Vazire & Oltmanns, 2013).

Second, we recommend that clinicians pose hypotheses about what patterns among informants' reports may reflect and test these hypotheses through focused follow-up assessment. By taking a hypothesis-testing approach to collecting informants' reports, the process of detecting patterns of discrepant and converging reports becomes a tool to facilitate accurate clinical decision-making, rather than a hindrance to sound decision-making. Essentially, we advocate for an approach that shares features with the comprehensive assessment approaches used in research on multi-informant assessments. For example, consider a case where a clinician is assessing a child client's aggressive behavior using parent and teacher reports, and finds that the teacher reports clinically significant aggression, whereas the parent reports minimal aggression. This presents an opportunity for the clinician to identify antecedents and triggers across home and school contexts that may be leading to contextual differences in the child's behavior. The clinician can test hypotheses about context-specific aggressive behavior through observing the child in parent-child interactions as well as the classroom environment. As another example, consider a case where a clinician is assessing an adult client's avoidant and anxious behaviors and finds that their client reports minimal behaviors, whereas their spouse reports clinically significant behaviors. This represents an opportunity for the clinician to gather information about their client's insight as well as the impact of their behavior on important social relationships. The clinician can test hypotheses about the client's insight and impairment through role-play interactions, followed up by activities focused on processing the client's performance and providing feedback on the interactions. In both of these examples, the clinician gains a richer conceptualization of the client's concerns through using informant reports, and perhaps most importantly, identifies treatment targets that are specific to the real-world contexts in which the client encounters challenges. These comprise some of the very reasons why we collect

informants' reports, and thus we should use the discrepancies that would logically manifest from these reports as tools to enhance clinical decision-making.

14.0 Conclusion*

In this chapter, we provided an overview of a commonly used, efficient clinical assessment method—informants' subjective reports about clients' mental health—and described research and theory on this method. Informants' reports commonly disagree with one another, and these disagreements present practical challenges in all of the settings in which assessors collect these data. However, in recent years conceptual and measurement models have informed research on links between scores on informants' reports and scores on non-subjective methods (e.g., observed behavior, performance-based tasks, official records). This research and theory has informed strategies for integrating informants' reports and maximizing their utility in clinical research and practice. Further, the existing evidence base allowed us to make recommendations on use of informants' reports in practice settings. Yet, although assessors have the benefit of decades of research on use and interpretation of informants' reports, much is left to learn about these reports and their ability to inform clinical work. In particular, we encourage future research on integrative, multi-informant strategies that can be tailored to individual clients, as well as work focused on understanding links between contextual variations in mental health and the informant discrepancies observed in multi-informant assessments of adult mental health.

15.0 References

- Achenbach, T. M. (2006). As others see us: Clinical and research implications of cross-informant correlations for psychopathology. *Current Directions in Psychological Science* **15**, 94-98.
- Achenbach, T. M. (2017). Future directions for clinical research, services, and training: Evidence-based assessment across informants, cultures, and dimensional hierarchies. *Journal of Clinical Child and Adolescent Psychology* **46**, 159-169.
- Achenbach, T. M. (2020). Bottom-up and top-down paradigms for psychopathology: A half century odyssey. *Annual Review of Clinical Psychology* **16**, 1-24.
- Achenbach, T. M., Ivanova, M. Y., & Rescorla, L. A. (2017). Empirically based assessment and taxonomy of psychopathology for ages 1½–90+ years: Developmental, multi-informant, and multicultural findings. *Comprehensive Psychiatry* **79**, 4-18.
- Achenbach, T. M., Krukowski, R. A., Dumenci, L., & Ivanova, M. Y. (2005). Assessment of adult psychopathology: Meta-analyses and implications of cross-informant correlations. *Psychological Bulletin* **131**, 361-382.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin* **101**, 213-232.
- Achenbach, T. M., & Rescorla, L. (2001). *Manual for the ASEBA school-age forms and profiles*. University of Vermont: Research Center for Children, Youth, and Families.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.
- Ashar, Y. K., Chang, L. J., & Wager, T. D. (2017). Brain mechanisms of the placebo effect: An affective appraisal account. *Annual Review of Clinical Psychology* **13**, 73-98.
- Bauer, D. J., Howard, A. L., Baldasaro, R. E., Curran, P. J., Hussong, A. M., Chassin, L., & Zucker, R. A. (2013). A trifactor model for integrating ratings across multiple informants. *Psychological Methods* **18**, 475-493.
- Becker-Haimes, E. M., Tabachnick, A. R., Last, B. S., Stewart, R. E., Hasan-Granier, A., & Beidas, R. S. (2020). Evidence base update for brief, free, and accessible youth mental health measures. *Journal of Clinical Child & Adolescent Psychology* **49**, 1-17.
- Becker-Haimes, E. M., Jensen-Doss, A., Birmaher, B., Kendall, P. C., & Ginsburg, G. S. (2018). Parent–youth informant disagreement: Implications for youth anxiety treatment. *Clinical Child*

Psychology and Psychiatry **23**, 42-56.

Beidas, R. S., Stewart, R. E., Walsh, L., Lucas, S., Downey, M. M., Jackson, K., . . . Mandell, D. S. (2015). Free, brief, and validated: Standardized instruments for low-resource mental health settings. *Cognitive and Behavioral Practice* **22**, 5-19.

Beidel, D. C., Turner, S. M., & Morris, T. L. (1995). A new inventory to assess childhood social anxiety and phobia: The Social Phobia and Anxiety Inventory for Children. *Psychological Assessment* **7**, 73-79.

Bird, H. R., Gould, M. S., & Staghezza, B. (1992). Aggregating data from multiple informants in child psychiatry epidemiological research. *Journal of the American Academy of Child and Adolescent Psychiatry* **31**, 78-85.

Blais, M. A., Hilsenroth, M. J., Castlebury, F., Fowler, J. C., & Baity, M. R. (2001). Predicting *DSM-IV* cluster B personality disorder criteria from MMPI-2 and Rorschach data: A test of incremental validity. *Journal of Personality Assessment* **76**, 150-168.

Bögels, S. M., Alden, L., Beidel, D. C., Clark, L. A., Pine, D. S., Stein, M. B., & Voncken, M. (2010). Social anxiety disorder: questions and answers for the DSM-V. *Depression and Anxiety* **27**, 168-189.

Borsboom, D. (2005). *Measuring the mind*. New York: Cambridge University Press.

Bossuyt, P. M., Reitsma, J. B., Bruns, D. E., Gatsonis, C. A., Glasziou, P. P., Irwig, L. M., . . . Lijmer, J. G. (2003). The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Clinical Chemistry* **49**, 7-18.

Brett, J. F., & Atwater, L. E. (2001). 360° feedback: Accuracy, reactions, and perceptions of usefulness. *Journal of Applied Psychology* **86**, 930-942.

Brown-Jacobsen, A. M., Wallace, D. P., & Whiteside, S. P. H. (2011). Multimethod, multi-informant agreement, and positive predictive value in the identification of child anxiety disorders using the SCAS and ADIS-C. *Assessment* **18**, 382-392.

Cannon, C. J., Makol, B. A., Keeley, L. M., Qasmieh, N., Okuno, H., Racz, S.J., & De Los Reyes, A. (2020). A paradigm for understanding adolescent social anxiety with unfamiliar peers: Conceptual foundations and directions for future research. *Clinical Child and Family Psychology Review* **23**, 338-364.

Carlson, E. N., Vazire, S., & Oltmanns, T. F. (2013). Self-other knowledge asymmetries in personality pathology. *Journal of Personality* **81**, 155-170.

- Casey, R. J., & Berman, J. S. (1985). The outcomes of psychotherapy with children. *Psychological Bulletin* **98**, 388-400.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Erlbaum.
- Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Cicchetti, D. (1984). The emergence of developmental psychopathology. *Child Development* **55**, 1-7.
- Cruitt, P. J., Hill, P. L., & Oltmanns, T. F. (in press). Personality pathology predicts increased informant-reported, but not performance-based, cognitive decline: Findings from two samples. *Personality Disorders: Theory, Research, and Treatment*. Advance online publication. <https://doi.org/10.1037/per0000434>
- Cuijpers, P., Andersson, G., Donker, T., & van Straten, A. (2011). Psychological treatment of depression: results of a series of meta-analyses. *Nordic Journal of Psychiatry* **65**, 354-364.
- Cuijpers, P., Li, J., Hofmann, S. G., & Andersson, G. (2010). Self-reported versus clinician-rated symptoms of depression as outcome measures in psychotherapy research on depression: a meta-analysis. *Clinical Psychology Review* **30**, 768-778.
- Cuijpers, P., Sijbrandij, M., Koole, S., Huibers, M., Berking, M., & Andersson, G. (2014). Psychological treatment of generalized anxiety disorder: a meta-analysis. *Clinical Psychology Review* **34**, 130-140.
- Cuming, S., Rapee, R. M., Kemp, N., Abbott, M. J., Peters, L., & Gaston, J. E. (2009). A self-report measure of subtle avoidance and safety behaviors relevant to social anxiety: development and psychometric properties. *Journal of Anxiety Disorders* **23**, 879-883.
- Dawes, R. M. (1999). Two methods for studying the incremental validity of a Rorschach variable. *Psychological Assessment* **11**, 297-302.
- De Clercq, B., De Fruyt, F., De Bolle, M., Van Hiel, A., Markon, K. E., & Krueger, R. F. (2014). The hierarchical structure and construct validity of the PID-5 trait measure in adolescence. *Journal of Personality* **82**, 158-169.
- De Los Reyes, A. (2011). More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child and Adolescent Psychology* **40**, 1-9.

De Los Reyes, A., Alfano, C. A., Lau, S., Augenstein, T. M., & Borelli, J. L. (2016a). Can we use convergence between caregiver reports of adolescent mental health to index severity of adolescent mental health concerns? *Journal of Child and Family Studies* **25**, 109-123.

De Los Reyes, A., Augenstein, T.M., & Aldao, A. (2017a). Assessment issues in child and adolescent psychotherapy. In J.R. Weisz and A.E. Kazdin (Eds.), *Evidence-based psychotherapies for children and adolescents* (3rd ed., pp. 537-554). New York: Guilford.

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D.A.G., Burgers, D., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin* **141**, 858-900.

De Los Reyes, A., Bunnell, B. E., & Beidel, D. C. (2013b). Informant discrepancies in adult social anxiety disorder assessments: Links with contextual variations in observed behavior. *Journal of Abnormal Psychology* **122**, 376-386.

De Los Reyes, A., Cook, C. R., Gresham, F. M., Makol, B. A., & Wang, M. (2019c). Informant discrepancies in assessments of psychosocial functioning in school-based services and research: Review and directions for future research. *Journal of School Psychology* **74**, 74-89.

De Los Reyes, A., Drabick, D. A. G., Makol, B. A., & Jakubovic, R. (2020). Introduction to the special section: The Research Domain Criteria's units of analysis and cross-unit correspondence in youth mental health research. *Journal of Clinical Child and Adolescent Psychology* **49**, 279-296.

De Los Reyes, A., Henry, D. B., Tolan, P. H., & Wakschlag, L. S. (2009). Linking informant discrepancies to observed variations in young children's disruptive behavior. *Journal of Abnormal Child Psychology* **37**, 637-652.

De Los Reyes, A., & Kazdin, A. E. (2005). Informant discrepancies in the assessment of childhood psychopathology: A critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin* **131**, 483-509.

De Los Reyes, A., & Kazdin, A. E. (2006). Conceptualizing changes in behavior in intervention research: The range of possible changes model. *Psychological Review* **113**, 554-583.

De Los Reyes, A., & Kazdin, A. E. (2008). When the evidence says, "Yes, no, and maybe so": Attending to and interpreting inconsistent findings among evidence-based interventions. *Current Directions in Psychological Science* **17**, 47-51.

De Los Reyes, A., & Kazdin, A. E. (2009). Identifying evidence-based interventions for children

and adolescents using the range of possible changes model: A meta-analytic illustration. *Behavior Modification* **33**, 583- 617.

De Los Reyes, A., Kundey, S.M.A., & Wang, M. (2011). The end of the primary outcome measure: A research agenda for constructing its replacement. *Clinical Psychology Review* **31**, 829-838.

De Los Reyes, A., & Langer, D. A. (2018). Assessment and the *Journal of Clinical Child and Adolescent Psychology*'s Evidence Base Updates series: Evaluating the tools for gathering evidence. *Journal of Clinical Child and Adolescent Psychology* **47**, 357-365.

De Los Reyes, A., Lerner, M. D., Keeley, L. M., Weber, R., Drabick, D. A. G., Rabinowitz, J., & Goodman, K. L. (2019a). Improving interpretability of subjective assessments about psychological phenomena: A review and cross-cultural meta-analysis. *Review of General Psychology* **23**, 293-319.

De Los Reyes, A., & Makol, B. A. (in press). Interpreting convergences and divergences in multi-informant, multi-method assessment. In J. Mihura (Ed.), *The Oxford handbook of personality and psychopathology assessment*. (2nd ed.). New York: Oxford.

De Los Reyes, A., Makol, B. A., Racz, S. J., Youngstrom, E. A., Lerner, M. D., & Keeley, L. M. (2019b). The Work and Social Adjustment Scale for Youth: A measure for assessing youth psychosocial impairment regardless of mental health status. *Journal of Child and Family Studies* **28**, 1-16.

De Los Reyes, A., & Ohannessian, C.M., (2016). Introduction to the special issue: Discrepancies in adolescent-parent perceptions of the family and adolescent adjustment. *Journal of Youth and Adolescence* **45**, 1957-1972.

De Los Reyes, A., Ohannessian, C. M., & Laird, R. D. (2016b). Developmental changes in discrepancies between adolescents' and their mothers' views of family communication. *Journal of Child and Family Studies* **25**, 790-797.

De Los Reyes, A., Thomas, S. A., Goodman, K. L., & Kundey, S. M. A. (2013a). Principles underlying the use of multiple informants' reports. *Annual Review of Clinical Psychology* **9**, 123-149.

Deros, D. E., Racz, S. J., Lipton, M. F., Augenstein, T. M., Karp, J. N., Keeley, L. M., Qasmieh, N., Grewe, B., Aldao, A., & De Los Reyes, A. (2018). Multi-informant assessments of adolescent

social anxiety: Adding clarity by leveraging reports from unfamiliar peer confederates. *Behavior Therapy* **49**, 84-98.

Di Nardo, P. A., Brown, T. A., & Barlow, D. H. (1995). *Anxiety Disorders Interview Schedule for DSM-IV* (lifetime version). San Antonio, TX: Psychological Corporation.

Dirks, M. A., Boyle, M. H., & Georgiades, K. (2011). Psychological symptoms in youth and later socioeconomic functioning: Do associations vary by informant? *Journal of Clinical Child and Adolescent Psychology* **40**, 10-22.

Dirks, M. A., De Los Reyes, A., Briggs-Gowan, M. J., Cella, D., & Wakschlag, L. S. (2012). Embracing not erasing contextual variability in children's behavior—theory and utility in the selection and use of methods and informants in developmental psychopathology. *Journal of Child Psychology and Psychiatry* **53**, 558-574.

Drabick, D. A. G., Gadow, K. D., & Loney, J. (2007). Source-specific oppositional defiant disorder: Comorbidity and risk factors in referred elementary schoolboys. *Journal of the American Academy of Child and Adolescent Psychiatry* **46**, 92-101.

Drabick, D. A. G., Gadow, K. D., & Loney, J. (2008). Co-occurring ODD and GAD symptom groups: Source-specific syndromes and cross-informant comorbidity. *Journal of Clinical Child and Adolescent Psychology* **37**, 314-326.

Frick, P. J. (1991). *The Alabama Parenting Questionnaire*. Unpublished rating scale, University of Alabama.

Frick, P. J., Christian, R. E., Wootton, J. M. (1999). Age trends in association between parenting practices and conduct problems. *Behavior Modification* **23**, 106-128.

Garb, H. N. (2003). Incremental validity and the assessment of psychopathology in adults. *Psychological Assessment* **15**, 508-520.

Goodman, S. H., & Gotlib, I. H. (1999). Risk for psychopathology in the children of depressed mothers: A developmental model for understanding mechanisms of transmission. *Psychological Review* **106**, 458-490.

Hartley, A. G., Zakriski, A. L., & Wright, J. C. (2011). Probing the depths of informant discrepancies: Contextual influences on divergence and convergence. *Journal of Clinical Child and Adolescent Psychology* **40**, 54-66.

- Hawley, K. M., & Weisz, J. R. (2003). Child, parent, and therapist (dis)agreement on target problems in outpatient therapy: The therapist's dilemma and its implications. *Journal of Consulting and Clinical Psychology* **71**, 62-70.
- Haynes, S. N., & Lench, H. C. (2003). Incremental validity of new clinical assessment measures. *Psychological Assessment* **15**, 456-466.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature* **466**, 29-29.
- Holmbeck, G. N., Li, S. T., Schurman, J. V., Friedman, D., & Coakley, R. M. (2002). Collecting and managing multisource and multimethod data in studies of pediatric populations. *Journal of Pediatric Psychology* **27**, 5-18.
- Hou, Y., Benner, A. D., Kim, S. Y., Chen, S., Spitz, S., Shi, Y., & Beretvas, T. (2019). Discordance in parents' and adolescents' reports of parenting: A meta-analysis and qualitative review. *American Psychologist* **75**, 329-348.
- Hunsley, J., & Lee, C. M. (2014). *Introduction to clinical psychology* (2nd ed.). Hoboken, NJ: Wiley.
- Hunsley, J., & Mash, E. J. (2007). Evidence-based assessment. *Annual Review of Clinical Psychology* **3**, 29-51.
- Hunsley, J., & Mash, E. J. (Eds.). (2018). *A guide to assessments that work* (2nd ed.). New York, NY: Oxford University Press.
- Hunsley, J., & Mash, E. J. (Eds.). (2008). *A guide to assessments that work*. New York, NY: Oxford University Press.
- Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment* **15**, 446-455.
- Jensen, P. S., Rubio-Stipec, M., Canino, G., Bird, H. R., Dulcan, M. K., Schwab-Stone, M. E., et al. (1999). Parent and child contributions to diagnosis of mental disorder: Are both informants always necessary? *Journal of the American Academy of Child & Adolescent Psychiatry* **38**, 1569-1579.
- Jones, J. D., Boyd, R. C., Calkins, M. E., Ahmed, A., Moore, T. M., Barzilay, R., Benton, T. D., & Gur, R. E. (2019). Parent-adolescent agreement about adolescents' suicidal thoughts. *Pediatrics* **143**, 1-12.

- Jopp, A. M., & South, S. C. (2015). Investigating the Personality Inventory for DSM-5 using self and spouse reports. *Journal of Personality Disorders* **29**, 193-214.
- Jordan, A. K., Thomeer, M. L., Lopata, C., Donnelly, J. P., Rodgers, J. D., & McDonald, C. A. (2019). Informant discrepancies in the assessment of adaptive behavior of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders* **49**, 2024-2034.
- Kazdin, A. E. (2017). *Research design in clinical psychology* (5th ed.). Boston, MA: Pearson.
- Kazdin, A. E. (1989). Identifying depression in children: A comparison of alternative selection criteria. *Journal of Abnormal Child Psychology* **17**, 437-455.
- Kazdin, A. E., & Kagan, J. (1994). Models of dysfunction in developmental psychopathology. *Clinical Psychology: Science and Practice* **1**, 35-52.
- Kazdin, A. E., & Rotella, C. (2009). *The Kazdin method for parenting the defiant child: With no pills, no therapy, no contest of wills*. Houghton Mifflin Harcourt.
- Keeley, L. M., Makol, B. A., Qasmieh, N., Deros, D. E., Karp, J. N., Lipton, M. F. et al. (2018). Validity of adolescent and parent reports on the Six-Item ADHD Self-Report Scale (ASRS-6) in clinical assessments of adolescent social anxiety. *Journal of Child and Family Studies* **27**, 1041-1053.
- Kessler, R. C., Adler, L. A., Gruber, M. J., Sarawate, C. A., Spencer, T., & van Brunt, D. L. (2007). Validity of the World Health Organization Adult ADHD Self-Report Scale (ASRS) screener in a representative sample of health plan members. *International Journal of Methods in Psychiatric Research* **16**, 52-65.
- Korelitz, K. E., & Garber, J. (2016). Congruence of parents' and children's perceptions of parenting: A meta-analysis. *Journal of Youth and Adolescence* **45**, 1973-1995.
- Kovacs, M., & Preiss, M. (1992). *Children's Depression Inventory*. New York: Multi-Health Systems.
- Kraemer, H. C., Measelle, J., Ablow, J., Essex, M., Boyce, W. T., & Kupfer, D. (2003). A new approach to integrating data from multiple informants in psychiatric assessment and research: Mixing and matching contexts and perspectives. *American Journal of Psychiatry* **160**, 1566-1577.
- Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine* **42**, 1879-1890.

- Kwon, K., Kim, E. M., & Sheridan, S. M. (2012). A contextual approach to social skills assessment in the peer group: Who is the best judge? *School Psychology Quarterly* **27**, 121-133.
- Laird, R. D., & De Los Reyes, A. (2013). Testing informant discrepancies as predictors of early adolescent psychopathology: Why difference scores cannot tell you what you want to know and how polynomial regression may. *Journal of Abnormal Child Psychology* **41**, 1-14.
- Laird, R.D., & LaFleur, L.K. (2016). Disclosure and monitoring as predictors of mother–adolescent agreement in reports of early adolescent rule-breaking behavior. *Journal of Clinical Child and Adolescent Psychology* **45**, 188-200.
- Lambert, M. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research* **17**, 1-14.
- Lambert, M. J., Hatch, D. R., Kingston, M. D., & Edwards, B. C. (1986). Zung, Beck, and Hamilton Rating Scales as measures of treatment outcome: A meta-analytic comparison. *Journal of Consulting and Clinical Psychology* **54**, 54-59.
- Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta-analysis. *Clinical Psychology: Science and Practice* **10**, 288-301.
- Lapouse, R., & Monk, M. A. (1958). An epidemiologic study of behavior characteristics in children. *American Journal of Public Health* **48**, 1134-1144.
- Lerner, M. D., De Los Reyes, A., Drabick, D. A. G., Gerber, A. H., & Gadow, K. D. (2017). Informant discrepancies define discrete, clinically useful autism spectrum disorder subgroups. *Journal of Child Psychology and Psychiatry* **58**, 829-839.
- Lippold, M. A., Greenberg, M. T., & Collins, L. M. (2013). Parental knowledge and youth risky behavior: A person oriented approach. *Journal of Youth and Adolescence* **42**, 1732-1744.
- Lippold, M. A., Greenberg, M. T., & Collins, L. M. (2014). Youths' substance use and changes in parental knowledge-related behaviors during middle school: A person-oriented approach. *Journal of Youth and Adolescence* **43**, 729-744.
- Luthar, S. S., Cicchetti, D., & Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development* **71**, 543-562.
- MacLeod, R. J., McNamee, J. E., Boyle, M. H., Offord, D. R., & Friedrich, M. (1999). Identification of childhood psychiatric disorder by informant: Comparisons of clinic and community samples. *Canadian Journal of Psychiatry* **44**, 144-150.

- Makol, B. A., De Los Reyes, A., Garrido, E., Harlaar, N., & Taussig, H. (2021). Assessing the mental health of maltreated youth with child welfare involvement using multi-informant reports. *Child Psychiatry and Human Development* **52**, 49-62.
- Makol, B. A., De Los Reyes, A., Ostrander, R., & Reynolds, E. K. (2019). Parent-youth divergence (and convergence) in reports of youth internalizing problems in psychiatric inpatient care. *Journal of Abnormal Child Psychology* **47**, 1677-1689.
- Makol, B. A., & Polo, A. J. (2018). Parent-child endorsement discrepancies among youth at chronic-risk for depression. *Journal of Abnormal Child Psychology* **46**, 1977-1088.
- Makol, B. A., Youngstrom, E. A., Racz, S. J., Qasmieh, N., Glenn, L. E., & De Los Reyes, A. (2020). Integrating multiple informants' reports: How conceptual and measurement models may address long-standing problems in clinical decision-making. *Clinical Psychological Science* **8**, 953-970.
- Markon, K. E., Quilty, L. C., Bagby, R. M., & Krueger, R. F. (2013). The development and psychometric properties of an informant report form of the Personality Inventory for DSM-5 (PID-5). *Assessment* **20**, 370-383.
- Marsh, J. K., De Los Reyes, A., & Lilienfeld, S. O. (2018). Leveraging the multiple lenses of psychological science to inform clinical decision making: Introduction to the special section. *Clinical Psychological Science* **6**, 167-176.
- Marsh, J.K., Zeveney, A., & De Los Reyes, A. (2020). Informant discrepancies in judgments about change during mental health treatments. *Clinical Psychological Science* **8**, 318-332.
- McCutcheon, A.L. (1987). *Latent class analysis*. Sage.
- McClendon, D. T., Warren, J. S., M. Green, K., Burlingame, G. M., Eggett, D. L., & McClendon, R. J. (2011). Sensitivity to change of youth treatment outcome measures: A comparison of the CBCL, BASC-2, and Y-OQ. *Journal of Clinical Psychology* **67**, 111-125.
- Measelle, J. R., Ablow, J. C., Cowan, P. A., & Cowan, C. P. (1998). Assessing young children's views of their academic, social, and emotional lives: an evaluation of the self-perception scales of the Berkeley Puppet Interview. *Child Development* **69**, 1556-1576.
- Miller, L. D., Martinez, Y. J., Shumka, E., & Baker, H. (2014). Multiple informant agreement of child, parent, and teacher ratings of child anxiety within community samples. *The Canadian Journal of Psychiatry* **59**, 34-39.

- Mischel, W., & Shoda, Y. (1995). A cognitive-affective system theory of personality: Reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. *Psychological Review* **102**, 246-268.
- Müller, J. M., Romer, G., & Achtergarde, S. (2014). Correction of distortion in distressed mothers' ratings of their preschool-aged children's Internalizing and Externalizing scale score. *Psychiatry Research* **215**, 170-175.
- Mundt, J. C., Marks, I. M., Shear, M. K., & Greist, J. M. (2002). The Work and Social Adjustment Scale: a simple measure of impairment in functioning. *The British Journal of Psychiatry* **180**, 461-464.
- Munkvold, L., Lundervold, A., Lie, S. A., & Manger, T. (2009). Should there be separate parent and teacher-based categories of ODD? Evidence from a general population. *Journal of Child Psychology and Psychiatry* **50**, 1264-1272.
- Nelemans, S. A., Branje, S. J., Hale, W. W., Goossens, L., Koot, H. M., Oldehinkel, A. J., & Meeus, W. H. (2016). Discrepancies between perceptions of the parent-adolescent relationship and early adolescent depressive symptoms: An illustration of polynomial regression analysis. *Journal of Youth and Adolescence* **45**, 2049-2063.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw Hill.
- Offord, D. R., Boyle, M. H., Racine, Y., Szatmari, P., Fleming, J. E., Sanford, M., et al. (1996). Integrating assessment data from multiple informants. *Journal of the American Academy of Child and Adolescent Psychiatry* **35**, 1078-1085.
- Ogles, B. M., Lambert, M. J., Weight, D. G., & Payne, I. R. (1990). Agoraphobia outcome measurement: A review and meta-analysis. *Psychological Assessment* **2**, 317-325.
- Oltmanns, T. F., & Turkheimer, E. (2009). Person perception and personality pathology. *Current Directions in Psychological Science* **18**, 32-36.
- Paul, S. E., Winograd, R. P., & Oltmanns, T. F. (in press). Personality pathology and substance misuse in later life: Perspectives from interviewer-, self-, and informant-reports. *Journal of Psychopathology and Behavioral Assessment*. Advance online publication. <https://doi.org/10.1007/s10862-020-09862-z>
- Pelham, Jr., W. E., Fabiano, G. A., & Massetti, G. M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child*

and *Adolescent Psychology* **34**, 449-476.

Piacentini, J. C., Cohen, P., & Cohen, J. (1992). Combining discrepant information from multiple sources: Are complex algorithms better than simple ones? *Journal of Abnormal Child Psychology* **20**, 51-63.

Piccirillo, M. L., Dryman, M. T., & Heimberg, R. G. (2016). Safety behaviors in adults with social anxiety: review and future directions. *Behavior Therapy* **47**, 675-687.

Qasmieh, N., Makol, B. A., Augenstein, T. M., Lipton, M. F., Deros, D. E., Karp, J. N., ... De Los Reyes, A. (2018). A multi-informant approach to assessing safety behaviors among adolescents: Psychometric properties of the Subtle Avoidance Frequency Examination. *Journal of Child and Family Studies* **27**, 1830-1843.

Renk, K., & Phares, V. (2004). Cross-informant ratings of social competence in children and adolescents. *Clinical Psychology Review* **24**, 239-254.

Renner, F., Cuijpers, P., & Huibers, M. J. H. (2014). The effect of psychotherapy for depression on improvements in social functioning: a meta-analysis. *Psychological Medicine* **44**, 2913.

Rescorla, L. A., Achenbach, T. M., Ivanova, M. Y., Turner, L. V., Árnadóttir, H., Au, A., ... & Zasepa, E. (2016). Collateral reports and cross-informant agreement about adult psychopathology in 14 societies. *Journal of Psychopathology and Behavioral Assessment* **38**, 381-397.

Rescorla, L. A., Bochicchio, L., Achenbach, T. M., Ivanova, M. Y., Almqvist, F., Begovac, I., ... & Verhulst, F. C. (2014). Parent teacher agreement on children's problems in 21 societies. *Journal of Clinical Child & Adolescent Psychology* **43**, 627-642.

Rescorla, L. A., Ewing, G., Ivanova, M. Y., Aebi, M., Bilenberg, N., Dieleman, G. C., ... & Verhulst, F. C. (2017). Parent-adolescent cross-informant agreement in clinically referred samples: Findings from seven societies. *Journal of Clinical Child & Adolescent Psychology* **46**, 74-87.

Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., & Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research* **18**, 169-184.

Reynolds, C., & Kamphaus, R. (2004). *Behavior Assessment System for Children (2nd ed.)*. Circle Pines, MN: American Guidance Service.

Rezeppa, T., Okuno, H., Qasmieh, N., Racz, S.J., Borelli, J.L., & De Los Reyes, A. (2021). Unfamiliar untrained observers' ratings of adolescent safety behaviors within social interactions

with unfamiliar peer confederates. *Behavior Therapy* **52**, 564-576.

Richters, J. E. (1992). Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin* **112**, 485-499.

Romano, E., Weegar, K., Babchishin, L., & Saini, M. (2018). Cross-informant agreement on mental health outcomes in children with maltreatment histories: A systematic review. *Psychology of Violence* **8**, 19-30.

Sanislow, C. A., Pine, D. S., Quinn, K. J., Kozak, M. J., Garvey, M. A., Heiassen, R. K., . . .

Cuthbert, B. N. (2010). Developing constructs for psychopathology research: Research domain criteria. *Journal of Abnormal Psychology* **119**, 631-639.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children Version IV (NIMH DISC-IV): description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry* **39**, 28-38.

Shimokawa, K., Lambert, M. J., & Smart, D. W. (2010). Enhancing treatment outcome of patients at risk of treatment failure: Meta-analytic and mega-analytic review of a psychotherapy quality assurance system. *Journal of Consulting and Clinical Psychology* **78**, 298-311.

Silverman, W. K., & Nelles, W. B. (1988). The Anxiety Disorders Interview Schedule for Children. *Journal of the American Academy of Child and Adolescent Psychiatry* **27**, 772-778.

Sparrow, S. S., Balla, D., & Cicchetti, D. (1984). *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service.

Stratis, E. A., & Lecavalier, L. (2015). Informant agreement for youth with autism spectrum disorder or intellectual disability: A meta-analysis. *Journal of Autism and Developmental Disorders* **45**, 1026-1041.

Sulik, M. J., Blair, C., Greenberg, M., & Family Life Project Investigators (2017). Child conduct problems across home and school contexts: A person-centered approach. *Journal of Psychopathology and Behavioral Assessment* **39**, 46-57.

Talbott, E., De Los Reyes, A., Power, T., Michel, J., & Racz, S. J. (2021). A team-based collaborative care model for youth with attention deficit hyperactivity disorder in education and pediatric health care settings. *Journal of Emotional and Behavioral Disorders* **29**, 24-33.

Thompson, E., Kline, E., Reeves, G., Pitts, S. C., Bussell, K., & Schiffman, J. (2014). Using parent and youth reports from the Behavior Assessment System for Children, to identify

individuals at clinical high-risk for psychosis. *Schizophrenia Research* **154**, 107-112.

Turner, S. M., Beidel, D. C., Dancu, C. V., & Stanley, M. A. (1989). An empirically derived inventory to measure social fears and anxiety: the Social Phobia and Anxiety Inventory. *Psychological Assessment* **1**, 35-40.

van der Ende, J., Verhulst, F. C., & Tiemeier, H. (2012). Agreement of Informants on emotional and behavioral problems from childhood to adulthood. *Psychological Assessment* **24**, 293-300.

Vazire, S. (2006). Informant reports: A cheap, fast, and easy method for personality assessment. *Journal of Research in Personality* **40**, 472-481.

Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology* **98**, 281-300.

Vazire, S., & Mehl, M. R. (2008). Knowing me, knowing you: The accuracy and unique predictive validity of self-ratings and other-ratings of daily behavior. *Journal of Personality and Social Psychology* **95**, 1202-1216.

Weisz, J. R., Jensen Doss, A., & Hawley, K. M. (2005). Youth psychotherapy outcome research: A review and critique of the evidence base. *Annual Review of Psychology* **56**, 337-363.

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin* **132**, 132-149.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: A meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology* **55**, 542–549.

Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B., . . . Bossuyt, P. M. (2011). QUADAS-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* **155**, 529-536.

Xu, Y., Boyd, R. C., Butler, L., Moore, T. M., & Benton, T. D. (2017). Associations of parent adolescent discrepancies in family cohesion and conflict with adolescent impairment. *Journal of Child and Family Studies* **26**, 3360-3369.

Yeh, M., & Weisz, J. R. (2001). Why are we here at the clinic? Parent-child (dis)agreement on referral problems at outpatient treatment entry. *Journal of Consulting and Clinical Psychology* **69**, 1018-1025.

Youngstrom, E., Izard, C., & Ackerman, B. (1999). Dysphoria-related bias in maternal ratings of children. *Journal of Consulting and Clinical Psychology* **67**, 905-916.

Youngstrom, E. A., Findling, R. L., & Calabrese, J. R. (2003). Who are the comorbid adolescents? Agreement between psychiatric diagnosis, youth, parent, and teacher report. *Journal of Abnormal Child Psychology* **31**, 231-245.

Youngstrom, E. A., Findling, R. L., Calabrese, J. R., Gracious, B. L., Demeter, C., DelPorto Bedoya, D., & Price, M. (2004). Comparing the diagnostic accuracy of six potential screening instruments for bipolar disorder in youths aged 5 to 17 years. *Journal of the American Academy of Child and Adolescent Psychiatry* **43**, 847-858.

Youngstrom, E. A. (2008). Evidence-based strategies for the assessment of developmental psychopathology: Measuring prediction, prescription, and process. In D. J. Miklowitz, W. E. Craighead, & L. Craighead (Eds.), *Developmental psychopathology* (pp. 34–77).

New York, NY: Wiley.

Youngstrom, E. A., Genzlinger, J. E., Egerton, G. A., & Van Meter, A. R. (2015). Multivariate meta-analysis of the discriminative validity of caregiver, youth, and teacher rating scales for pediatric bipolar disorder: Mother knows best about mania. *Archives of Scientific Psychology* **3**, 112-137.

Youngstrom, E. A., Van Meter, A., Frazier, T. W., Hunsley, J., Prinstein, M. J., Ong, M. L., & Youngstrom, J. K. (2017). Evidence-based assessment as an integrative model for applying psychological science to guide the voyage of treatment. *Clinical Psychology: Science and Practice* **24**, 331-363.

Table 1
Examples of Multi-Informant Assessment Instruments

Measure	Domains Assessed	Age Groups	Items	Informants	Representative Citations
<i>Achenbach System of Empirically Based Assessment (ASEBA)</i>	Emotional and behavioral problems	1 ½ years or older	99-118	Parent-Report, Teacher-Report, Self-Report, Collateral-Report ^a	Achenbach & Rescorla, 2001; Achenbach et al., 2017; Rescorla et al., 2014, 2016, 2017
<i>Behavioral Assessment System for Children (BASC-2/BASC-3)</i>	Emotional and behavioral problems	2 to 21 years	150-176	Parent-Report, Teacher-Report, Self-Report	McClendon et al., 2011; Miller et al., 2014; Reynolds & Kamphaus, 2004; Thompson et al. 2014
<i>NEO-Personality Inventory–3 (NEO PI-3)</i>	Personality traits	17 years or older	240	Self-Report, Collateral-Report ^a	Costa & McCrae, 1992; Cruitt et al., in press; Paul et al., in press
<i>Personality Inventory for the DSM-5 (PID-5)</i>	Personality traits	11 years or older	220	Self-Report, Collateral-Report ^a	De Clercq et al., 2013; Jopp & South, 2015; Krueger et al., 2012; Markon et al., 2013
<i>Vineland Adaptive Behavior Scale (VABS)</i>	Adaptive behavior	Birth to 90 years	149-502	Parent-Report, Teacher-Report	Jordan et al., 2019; Sparrow et al., 1984

Note. ^a Category of informants most commonly used in adult assessments and who are nominated by the individual being evaluated. Common collateral informants include spouses, co-workers, and parents.

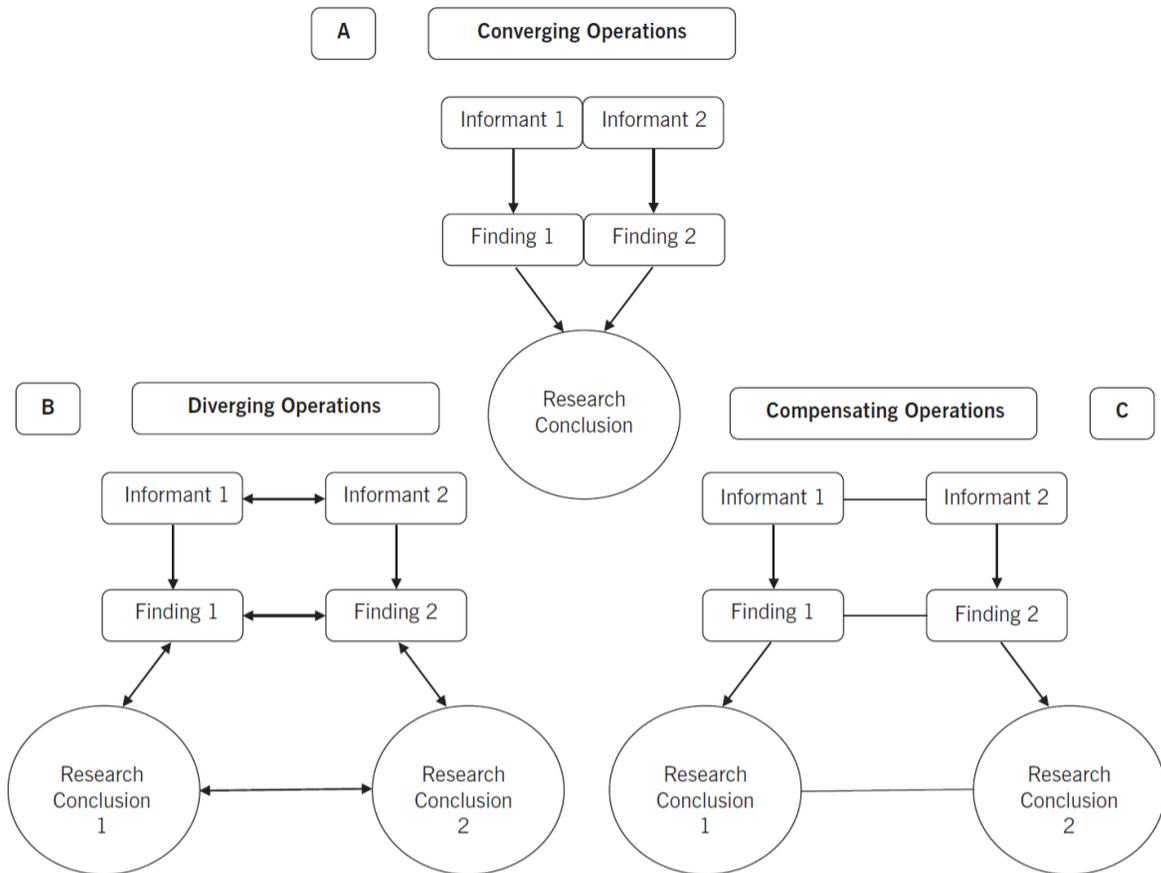


Figure 1. Graphical representation of the research concepts that comprise the Operations Triad Model. The top half (A) represents Converging Operations: a set of measurement conditions for interpreting patterns of findings based on the consistency within which findings yield similar conclusions. The bottom half denotes two circumstances within which researchers identify discrepancies across empirical findings derived from multiple informants' reports and thus discrepancies in the research conclusions drawn from these reports. On the left (B) is a graphical representation of Diverging Operations: a set of measurement conditions for interpreting patterns of inconsistent findings based on hypotheses about variations in the behavior(s) assessed. The solid lines linking informants' reports, empirical findings derived from these reports, and conclusions based on empirical findings denote the systematic relations among these three study components. Further, the presence of dual arrowheads in the figure representing Diverging Operations conveys the idea that one ties meaning to the discrepancies among empirical findings and research conclusions and thus how one interprets informants' reports to vary as a function of variation in the behaviors being assessed. Lastly, on the right (C) is a graphical representation of Compensating Operations: a set of measurement conditions for interpreting patterns of inconsistent findings based on methodological features of the study's measures or informants. The dashed lines denote the lack of systematic relations among informants' reports, empirical findings, and research conclusions. Originally published in De Los Reyes, Thomas, et al. (2013). © Annual Review of Clinical Psychology. Copyright 2012 Annual Reviews. All rights reserved. The Annual Reviews logo, and other Annual Reviews products referenced herein are either registered trademarks or trademarks of Annual Reviews. All other marks are the property of their respective owner and/or licensor.